SECU

AD-A218 268

...ENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. R
UNCLASSIFIED

1b. RESTRICTIVE MARKINGS
NONE

2a. SECURITY CLASSIFICATION AUTHORITY

3. DISTRIBUTION/AVAILABILITY OF REPORT
APPROVED FOR PUBLIC RELEASE;
UNLIMITED DISTRIBUTION

1

2b. DECLASSIFICATION / DOWNGRADING SCHEDULE

4. PERFORMING ORGANIZATION REPORT NUMBER(S)

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFIT/CI/CIA-89-128

6a. NAME OF PERFORMING ORGANIZATION
AFIT STUDENT AT UNIVERSITY OF COLORADO

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
AFIT/CIA

6c. ADDRESS (City, State, and ZIP Code)

7b. ADDRESS (City, State, and ZIP Code)
Wright-Patterson AFB OH 45433-6583

8a. NAME OF FUNDING/SPONSORING ORGANIZATION

8b. OFFICE SYMBOL (If applicable)

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER

8c. ADDRESS (City, State, and ZIP Code)

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification) (UNCLASSIFIED)
A Long-Term Retention Advantage for Spatial Information Learned Naturally and in the Laboratory

12. PERSONAL AUTHOR(S)
William T. Wittman

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
THESIS	FROM ______ TO ______	1989	190

16. SUPPLEMENTARY NOTATION
APPROVED FOR PUBLIC RELEASE IAW AFR 190-1
ERNEST A. HAYGOOD, 1st Lt, USAF
Executive Officer, Civilian Institution Programs

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

90 02 12 003

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
ERNEST A. HAYGOOD, 1st Lt, USAF

22b. TELEPHONE (Include Area Code)
513-255-2259

22c. OFFICE SYMBOL
AFIT/CI

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE

A LONG-TERM RETENTION ADVANTAGE FOR SPATIAL INFORMATION

LEARNED NATURALLY AND IN THE LABORATORY

by

WILLIAM THOMAS WITTMAN

B.A., California State University, Fresno, 1977

M.S., California State University, Fresno, 1979

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Psychology

1989

This thesis for the Doctor of Philosophy degree by

William Thomas Wittman

has been approved for the

Department of

Psychology

by

Alice F. Healy

Alice F. Healy

Walter Kintsch

Walter Kintsch

Date June 9, 1989

Wittman, William Thomas (Ph.D., Psychology)

A long-term retention advantage for spatial information learned naturally and in the laboratory

Thesis directed by Professor Alice F. Healy

The long-term retention characteristics of three memory components learned both naturally and in the laboratory were investigated. Using a cued recall procedure, 48 college students were asked to recall the spatial, temporal, and item components of their own semester schedules (Experiment 1), or a fictitious schedule (Experiment 2). In completing class schedule questionnaires, students were both cued with and asked to recall these three components. For example, a subject might be given the name of a course (item component) and then be asked to locate on a campus map where the class was held (spatial component). In Experiment 1, a longitudinal as well as cross-sectional approach was taken wherein subjects were tested three times, each time covering three different retention intervals. In total, recall data were gathered from intervals ranging from approximately 12 to 36 months in length. In Experiment 2, subjects studied a fictitious semester schedule and were tested approximately one week and six weeks following training. Results from both experiments indicated better retention of the spatial component of class schedules over either the item or temporal components.

All three components showed poorer recall over time, with the spatial component showing the greatest stability. Results were discussed in terms of a proceduralist view of memory performance (Kolers & Roediger, 1984). It was proposed that such accounts of performance must not ignore possible differences in the organization and representation of specific types of information in memory.

DEDICATION

To Mary and her Son

ACKNOWLEDGEMENTS

I wish to thank all the members of my dissertation committee, Dr. Alice Healy, Dr. Lyle Bourne, Dr. Walter Kintsch, Dr. Dave Thomas, and Dr. Verne Keenan, for their time, support, and technical advice. I especially want to thank Dr. Alice Healy for her patience and encouragement throughout this project.

I wish also to thank my parents, parents-in-law, and all my relatives and friends for their support and prayers.

Finally, and most importantly, I wish to thank my wife, Betty, and my children, Tommy, Molly, and Annie, for their love, continuous support, and for all the blessings they bring to my life.

This work was supported in part by the United States Army Research Institute Contract MDA903-86-K-0155 to the Institute of Cognitive Science at the University of Colorado.

CONTENTS

TABLES

Table

FIGURES

Figure

CHAPTER I

INTRODUCTION

Most of us at some time or another have forgotten the name of a street or a house address in an otherwise familiar town. While the information was not remembered, we may have felt certain, if given the opportunity, that we could physically locate the street or house. This example, like many others, suggests that people may be better at recalling certain types of information over others. The focus of this research is on just this question. That is, in the long term, is spatial information better recalled than temporal or item information. Any differences found have implications for our understanding of how knowledge is organized and represented in memory.

Attempts at delineating the types of knowledge stored in human memory abound in the cognitive research literature. Some of the better known divisions include the distinction between semantic, episodic and procedural knowledge (Tulving, 1972, 1985); between declarative and procedural knowledge (Fitts & Posner, 1967; Anderson, 1982, 1983); and between verbal and visual knowledge (Paivio, 1971, 1986). Unfortunately, the question of what is stored in memory has often been entangled with the question of what is presented for storage. The two need

not be the same. In fact, much of the controversy in this area has focused on how any one form of representation can accommodate the variety of seemingly dissimilar types of input.

The aim of this research is to separate the issues of representation and type of input. The objective is simply to examine whether certain types of knowledge or input lead to memory performance differences. If differences are found, only then should the possible types of representation be considered.

The basic division of knowledge used here comes from the work of Healy (1978, 1982), Lee and Estes (1981), and others on short-term memory, though the terms have been broadened to include a larger range of information. It is proposed that knowledge be studied in terms of its spatial, temporal and item components. The spatial component includes information about locations of objects, their spatial relations and distances (Evans & Pezdek, 1980; Golledge, Smith, Pellegrino, Doherty, & Marshall, 1985), as well as knowledge of how to proceed through space (Thorndyke & Hayes-Roth, 1982). Temporal information includes knowledge of dates and times (White, 1982), and the relative order of events (Healy, 1974). Item information includes verbal information, such as facts, figures, and names (i.e., Bahrick, Bahrick, & Wittlinger, 1975).

There is no attempt here to propose separate memory systems or that these three components of knowledge are the only ones possible. Nor is it intended that the traditional divisions of

knowledge cited above be replaced with this division. Instead, it is believed that by first looking at these basic components of what is presented for study, a better understanding of what is retained in memory can be gained.

Any distinction made between the types of knowledge is useful only if it is tied to actual differences in performance. Research providing evidence for the differential processing of spatial, temporal, and item information, and more specifically, the retention advantage for spatial information will be reviewed in two groups. The first group includes laboratory studies involving serial position functions in short-term memory, and a collection of other laboratory studies dealing specifically with spatial memory. The second includes studies of natural memory, that is, research examining learning accomplished in the real world, not in the laboratory.

Some of the most direct evidence for differences in the processing and retention of spatial, temporal, and item information comes from the short-term memory research dealing with the serial position effect. Early studies have suggested processing differences for item and order information in serial order recall. The basic notion is that one can remember a particular item from a list of say, letters, yet not recall its relative position in the list. Conversely, at times one can remember something was in, for example, the third serial position of a list yet not be able to name the item. For example, Healy

(1974) found that the serial position functions for consonants recalled in order were much more bowed than for recall of items with order constrained. This observation suggested that differences in the amount of bowing represented a loss of information not needed in simple item recall and that recall of item and order (temporal) information showed some degree of independence. Similar suggestions have been provided by Murdock (1976) and Shiffrin and Cook (1978).

Healy (1978) examined memory for item, order, and spatial information. As in earlier studies, subjects were shown four consonants randomly arranged in a linear array. This time however, order information was divided into recall of either the temporal sequence of the letters or their spatial locations in the array. Healy found that the serial position functions for spatial location were less bowed than those functions for temporal sequence information. Furthermore, spatial information showed a flatter retention function than temporal information. Healy (1982) further tested the temporal, spatial, item distinction, confirming their processing independence and suggesting that spatial information was retained longer and involved a different encoding strategy than temporal information.

A number of other laboratory studies of spatial memory also provide evidence for differential processing of spatial information. Salthouse (1974, 1975) in a short-term memory experiment had subjects recall either the spatial positions or

the identities of letters in a 25-letter, diamond-shaped array. Prior to recall subjects were given various intervening tasks designed to interfere selectively with remembering either letter identity or position. For example, subjects performed a same-different judgment task involving either faces or words. Memory performance was found to vary with the type of intervening task; interference was greatest if both tasks involved the same memory code. Since performance was related to the assumed memory code and the demands of the intervening tasks it was concluded that verbal and spatial information are stored and processed in separate information processing systems. In another line of research, Dean and Kulhavy (1981) found that subjects who read a 2190-word passage about a fictitious African tribe and drew a map related to the passage recalled the passage better than those not drawing a map. Similarly, studying a map prior to reading a passage facilitated recall. They concluded that use of a spatial organizer aided recall. Schwartz and Kulhavy (1981) examined more precisely the relationship of map features and text recall. Varying the characteristics of the map, they concluded that the spatial arrangement of the map features was the critical factor. Further evidence for the role of spatial information comes from an experiment by Pezdek, Roman, and Sobolik (1986). In their work they examined recall for two types of stimuli, 16 common objects and 16 one-word labels for these objects. Both objects and words were studied on a 6 X 6 matrix, arranged randomly.

Subjects were tested for recall of the items (names of objects or the words) and for the location of the items (using the matrix to place correctly the actual objects or word labels) after delays up to 90 seconds. They found that more objects were recalled and relocated than words. They attributed this advantage for objects to the encoding of spatial location information in study of the objects.

Natural memory research also provides evidence for differential processing of information and in some cases, for a spatial information advantage. Unfortunately, many natural memory studies have been more concerned with memory for when events have occurred, the temporal component, rather than either the spatial or item components. Fortunately, there are some exceptions to this trend. For example, several flashbulb memory studies have focused on specific classes of information evident in natural event memories (Brown & Kulik, 1977; Pillemer, 1984). These studies report that subjects followed a canonical form in describing their memories of critical events. This form included, among other categories, the where, what, and who of events, roughly equivalent to spatial and item information. These classes of information were consistently found in subject accounts of their memories and thus all were considered highly recalled, though no specific comparisons were reported. In another flashbulb memory study, Yarmey and Bull (1978) more directly studied differences in the components of natural

memories. They specifically questioned subjects about the where (spatial), when (temporal), who and what (item components) of their personal circumstances surrounding the assassination of President Kennedy. While actual differences in recall for these types of information were not reported, they did find that subject ratings of the clarity of each memory component to be about equal. In another line of research, Wagenaar (1978) studied memory for the component parts of radio traffic reports. He found that subjects remembered where traffic jams were located better than the names of the city or roads involved (item components). Wagenaar (1986) studied the long-term retention of the what, who, where, and when aspects of his own memories over a period of six years. Unfortunately, Wagenaar's (1986) interest was in the effectiveness of these types of information as cues for recall and not their relative retention characteristics. Retention data by aspect type were not discussed. Further natural memory evidence for differences in component processing is found in Bahrick, et al.'s (1975) study of long-term recognition memory for the names and faces of high school classmates. Among other memory tasks, subjects were shown either five pictures or five names of previous classmates. Subjects did far better in picture recognition than in name recognition. It was concluded that visual information (faces) was retained virtually unimpaired for 35 years while retention of verbal information (names) declined after 15 years. Thompson (1982), in

a diary study of memory for daily events, found that forgetting of when an event occurred (temporal component) and what occurred (item component) were related for the first few weeks following the event, but thereafter forgetting occurred at different rates. Temporal information followed a linear rate of forgetting while item information loss conformed to the more familiar negatively accelerated retention function.

In sum, the research literature from the laboratory provides evidence for processing differences among the spatial, temporal, and item components of memories. It suggests, at least in the short-term, there are retention differences across these three types of information. More specifically, memories involving spatial information are recalled better over time than those involving temporal or item information. While component differences, in some cases suggesting better retention of spatial information, are found in the natural memory literature, drawing any firm conclusions is difficult. Even though each of the studies cited deals with naturally learned memories, they represent a wide range of methods and materials. Furthermore, few of the studies was specifically designed to examine retention differences among memory components.

The goal of the current research was to examine whether the suggested retention advantage of spatial information found in the laboratory could be found using longer retention intervals, specifically, periods measured in months or years. Two

experiments were performed. The first experiment studied naturally learned experiences, those not learned in the laboratory. By using both a cross-sectional and longitudinal approach (see Bahrick, et al., 1975), memories were examined over a period of years without having to train subjects in the laboratory and then wait for forgetting to occur. Unfortunately, such natural memory studies do not allow the tight control of extraneous factors that is possible in the laboratory. Consequently, the present research looked at memory for natural events using a methodology that overcame many of the typical problems of such studies. It applied the advantages and reduces the disadvantages of three common approaches to the study of natural memory: the probe, diary and questionnaire methodologies.

Probe Methodology

In the probe method, subjects are provided with a probe or cue word, typically a common noun, and are asked to recall some associated life experience (see Crovitz & Schiffman, 1974; Franklin & Holding, 1977; Rubin, 1982). Once an experience is recalled, subjects are then asked to date their memories. Analysis typically focuses on the number of memories recalled over time. Strengths of the probe method include its simple, unstructured format and the ability to explore the effects of different types of retrieval cues or probes. On the other hand,

the approach is strongly criticized because the accuracy of subject self-reports cannot be determined.

Diary Methodology

In contrast to the probe method, the diary method allows the experimenter greater control over the experimental situation. Not only does the experimenter have control over the conditions of recall, but also, though to a lesser extent, the conditions of learning. In the typical diary method study, a subject is asked to record on a daily basis a description of one or more events or experiences that have occurred that day. At some point later, often a year or more, the subject is cued with each original event description and then asked to recall details of the event, often the date of the event (see Linton, 1978, 1982; Thompson, 1982; and White, 1982). While the diary method provides the experimenter with substantial control over what is remembered and how it is recalled, it is not without its criticisms. Because long-term diary keeping is tedious at best, these studies often involve single subject designs with the experimenter as subject. Beyond this limitation, all diary studies face the possible effects of bias in event selection and event recording as well as the effects on memory of rating events, for example, for their pleasantness.

Questionnaire Methodology

This third general approach to the study of natural memory

is similar to the probe method in that it involves no control of the learning of natural memories. It differs from the probe method by being more structured and more focused on specific memory episodes. In general, questionnaire studies simply involve formulating specific questions for subjects to answer about their everyday experiences. Such questionnaires are relatively inexpensive, easy to administer, and can be used with large numbers of subjects (see Brown & Kulik, 1977; Herrmann & Neisser, 1978; and Loftus & Marburger, 1983). A major criticism of the questionnaire method is the difficulty in assessing what the subject originally learned and to what extent it was learned.

Experiment 1 was designed to use the best aspects of each of these methods while avoiding their limitations. Like the probe method, the current study used cues to elicit recall of past events. Unlike the probe method though, recall focused on a specific time period in the past and cues applied to only one event during that period. Similar to the diary method, events to be recalled were specified prior to the experiment. Unlike the diary method, however, events were determined by the experimenter, not by the subject. Finally, like the questionnaire method, this study used a structured, question-then-response format for presentation of recall cues. Unlike most questionnaire studies however, there was a measure of what was originally learned. Beyond these improvements over typical methodologies this research incorporated a second

experiment designed specifically to test results from the real world in the laboratory. Details of Experiment 2 are provided later in the dissertation. The focus of the remainder of this chapter is on Experiment 1.

Experiment 1 used a cued recall format similar to that used by Wagenaar (1986). Wagenaar, in a natural memory study using the diary methodology, had his subject (himself), recall four aspects of his memories, the "who," "what," "where" and "when" of daily events. Wagenaar recorded these four aspects of his daily experiences for six years. At recall, Wagenaar cued himself with one or more of these aspects. He found that events cued with the "what" aspect showed better recall than those cued with the "when" aspect. Unfortunately, Wagenaar did not present an analysis of recall by aspect type. Visual inspection of his reported data suggest, at the very least, that "when" information was more poorly recalled than the "who," "what," or "where."

Like Wagenaar, the current study addressed the "who," "what," "where," and "when" of events, but avoided the disadvantages of the diary methodology. In keeping with the temporal, spatial, and item distinction made earlier, this study considered the "who" and "what" aspects as item information, whereas the "where" and "when" aspects were considered as spatial and temporal information, respectively. More specifically, Experiment 1 looked at students' memory for their course schedules learned naturally. College students were asked

questions about courses they took during previous semesters. Questions explored memory using the what, who, where, and when aspects used by Wagenaar (1986), but tailored to course schedules. After pilot testing a number of specific questions, four were selected that were not on the ceiling or floor. These included: Memory for the name of the course (what); for the name of the instructor (who); for the location of the course (where); and for the time the course took place (when). These four aspects are hereafter referred to as types of information. The advantage this study had over Wagenaar's approach was that answers to these questions were established prior to attempted recall, without the help or knowledge of subjects. As in Wagenaar's study, a cued recall paradigm was used wherein subjects were both asked to recall and were cued with the four types of information. By asking subjects to recall course information from three different semesters and after three retention intervals, recall was measured with delays ranging from approximately 12 to 36 months in length, in six month intervals. In addition to examining recall performance over time, a number of other factors were investigated. These included prior course experience, ratings of course enjoyment, gender differences, and the effectiveness of each type of information as a recall cue.

Overall, recall performance was predicted to be high. Performance will be strongly affected, at least for the first year, by the numerous rehearsals of the schedule information (two

to three times per week for a full semester). More importantly, following Healy (1978, 1982), where information, because it involved spatial information, was expected to produce the best performance. When information was predicted to be most poorly recalled and was expected to be a poor recall cue (Wagenaar, 1986). Also following Wagenaar's results, what information was expected to be the best recall cue, due to its relative uniqueness. In addition, positive correlations were expected between course experience and enjoyment ratings and recall performance. Both White (1982) and Wagenaar (1986) found pleasant events to be better recalled. Finally, gender differences were also examined, although it was uncertain whether male or female subjects would do better. Evidence has been mixed on this account as well. Robinson (1976) found clear gender differences using a prompt word methodology. Female subjects recalled more recent events than male subjects when given noun or verb prompt words. This difference was not found with adjectives prompts. The next chapter provides a complete discussion of the methods and the results of Experiment 1.

CHAPTER II

EXPERIMENT 1

The purpose of the present study was to examine the long-term retention characteristics of the spatial, temporal, and item components of memories learned naturally. To accomplish this goal, student memories for their semester schedules were selected for study. In this case, subjects learned, over the course of a semester, spatial information (where their classes were located), temporal information (class times), and two types of item information (the names of professors and class titles). These four kinds of information were examined both cross-sectionally and longitudinally resulting in five retention intervals ranging from 12 to 36 months, in 6 month increments. Memory performance was measured using a cued recall methodology which allowed accurate assessment of subject responses, thus overcoming many of the typical methodological problems of natural memory studies. It was predicted that the spatial component of natural memories would be better retained than either the temporal or item components.

Method

Subjects

Forty-eight University of Colorado Introductory Psychology students were used in this experiment, 29 male and 19 female subjects. All subjects had attended the University of Colorado for at least two years prior to initial testing. Following initial testing in Fall 1987, subjects were given course credit for participation. After participation in sessions 2 and 3, subjects were paid five dollars for each test day. All subjects received initial testing based on random assignment to one of three retention groups according to the semester being tested. One third (16) of the subjects were tested on courses taken during the semester approximately 12 months prior (Fall 1986). Likewise, the remaining two groups of 16 were tested on courses taken during the semester 18 and 24 months prior, Spring 1986 and Fall 1985, respectively. With testing sessions 2 and 3, subjects were again assigned to one of the three tested semesters such that after the last testing each subject had been tested on all three semesters. In addition, with each testing session the retention interval increased by six months for each semester being tested. With the second and third test sessions, there were 39 and 28 subjects participating, respectively, out of the original 48. Two of the subjects in the third session had not participated in the second session. Most of the nonparticipants

simply did not want to participate (they were no longer enrolled in an introductory psychology course). The remainder had moved or graduated. Specific subtotals of subjects participating in each condition are provided in the results section below.

Materials

Subjects were asked to complete a questionnaire covering three courses randomly selected from their tested semester schedule. The questionnaire was divided into three parts which cover information cues and questions, course experience, and semester experience, respectively.

In the first part of the questionnaire, subjects were cued with one of four types of course information, the what, who, where, or when of a course. Subjects were then asked to provide answers to questions using the remaining three types of course information. For example, subjects were cued with when a course took place (the class start time) and then asked to provide who instructed the course, where the classroom was located (on a campus map), and what the name of the course was. This was followed on the next page of the questionnaire with subjects getting two cues (for the same course), for example, the when cue again along with the who cue, with subjects completing the what and where information. This same cue-then-question procedure was repeated once for each of the three courses being tested.

Cues and questions were formed in the following manner. The

four types of information were grouped such that all 12 possible combinations were used, given that the same type of information was never grouped with itself. Thus, with each of the 12 combinations the first member of a pair served as the primary cue and the second served as the secondary cue.

Cue combinations were assigned to subjects according to the following criteria. In the three cue pairs assigned to each subject, the same primary cue type was never repeated and the same secondary cue was never repeated. Thus, all four information types were used at least once for each subject (either in the first or second position). Since each subject was shown all three pairs, the full cycle of 12 cue pairs were repeated with every four subjects. Two different sets of 12 cue pairs was constructed, each meeting the above subject criteria. In sum, these two sets were repeated six times to account for all forty-eight subjects in the study. On the second testing, subjects received the same three cue pairs used in the first testing, but each pair was not necessarily matched with the same course.

When answering a where question, subjects responded by marking the location of the classroom on a two-dimensional, black and white map of the University campus. The map included trees, shadowing (gave a 3-D appearance), and outlines of buildings (but no building or street names). The map did provide the direction of North and indicated the location of the Rocky Mountains. In

addition, when given a where cue, the subject was shown the class location on the map by an obvious mark next to the appropriate building. Appendix A provides a sample page from the questionnaire which includes the map.

In the second part of the questionnaire, subjects were given the correct answers to all four types of questions and then were asked to make several ratings about their experience with that course or like courses. Specifically, subjects were asked to rate how well they liked each of the four aspects of each course (on a scale from 1 to 7) and to rate their subsequent experience with each of the four aspects (on a scale from 0 to 4 or more). Appendix B provides the questions asked in this part of the questionnaire. As in the first part, these questions were repeated for each of the three courses being tested.

In the final section of the questionnaire, after answering questions for each course, subjects were asked to rate their liking for the entire semester, to determine how much they have reviewed their semester schedule since its completion, and to provide other biographical information. This part of the questionnaire was repeated only once, at the conclusion of testing.

Procedure

Prior to testing, subjects were gathered for a presession to obtain their consent to access their university academic records. Access to records was necessary to obtain the recall cues used in the questionnaire and to evaluate responses once the questionnaires were completed. Prior to being asked for consent, subjects completed a short survey about classes taken during the immediately preceding semester (not one of the actual test semesters). This task was intended only as a time-filler. The questions in the survey were not the same as those used in the actual test questionnaire. Subjects were told that their answers to the questions would be checked against their actual academic record for that semester. Only one subject did not grant consent. This subject was dismissed from the study, given credit for attending the presession, and replaced with another subject. Once consent was obtained, subjects scheduled for their initial testing session one week later. Before being dismissed subjects were told only that by the return session their answers to the survey questions would be evaluated. They were not told that a new questionnaire would also be administered.

In the interim, questionnaires were constructed, tailored to each subject's initial semester schedule. That is, three courses were identified for testing such that the what, who, where, and when of a course was not the same for any two courses. (Note: this standard was possible for all but one subject in the initial

testing, five subjects in session 2, and three subjects in session 3. For these exceptions subjects had two courses in the same building).

During the initial test session, subjects completed a questionnaire which matched their assigned semester, either Fall 1985, Spring 1986, or Fall 1986. Subjects were told they were being asked to recall information about courses taken during one of their previous semesters at the University. They were instructed to take careful note of the semester being tested before starting. They were asked to answer questions as accurately as possible, proceeding one page at a time without turning back to change answers or to note previous answers. Subjects were tested in small groups, usually not more than four at a time. Sessions lasted approximately 20 minutes.

Since this study measured retention of course information longitudinally as well as cross-sectionally, subjects were asked to return for two subsequent testings, once in the Spring 1988 semester and then again in the Fall 1988 semester. No presessions were necessary in these subsequent testings. Questionnaires were simply prepared using information previously obtained. Test session procedures were the same as in the initial testing. During the two return sessions, subjects were asked to recall courses for the two semesters not previously tested. For example, if subjects were tested on Fall 1986 courses during the initial session, they would be randomly

assigned testing on either the Spring 1986 or Fall 1985 semester courses during the second session and tested on the remaining semester during the third session. Thus, using a combined cross-sectional and longitudinal approach allowed the study of five retention periods, in six month increments, ranging from 12 to 30 months in length.

Results

Questionnaires were scored for the percentage of correct responses. Scoring was strict, with emphasis placed on consistent scoring across information types. A what response (course title) received a score of 1 if it contained all the words found in the title. Responses adding or missing function words, however, like "to" or, "of," were given full credit. Adding or leaving out content words resulted in only half credit (e.g., Beginning Russian 1 changed to Intro to Russian). A who response receiving full credit required the correct identification and spelling of an instructor's name. Accurate identifications, but misspelled names received half credit (i.e., Berbernes instead of Bebernes). Locations on the campus map (where responses) were given full credit if they precisely marked the correct building. Responses marking buildings immediately adjacent to the correct building were given half credit. A when response received full credit if it identified the precise time a course started. Half credit was given if the indicated start

time was within 30 minutes either side of the correct time.

The major interest in this analysis was whether information about where a course was held was retained better than information about course title, professor, or course time (the what, who, or when). Most of the analyses therefore focused on performance differences across these four information types. Of particular interest were differences among information types over time. Several other factors were also examined to determine their relationship to information type performance. These included degree of prior experience, ratings of enjoyment, gender, and cue effectiveness. All statistical tests in this experiment used a .05 level of confidence.

Information Type Differences and the Retention Interval

In the following analyses, a two-way analysis of variance procedure was used with information type as the single within-subjects factor and either semester or test session as the single between-subjects factor. The variable semester indicated the semester being tested, either Fall 1986, Spring 1986, or Fall 1985. At test session 1, these three semesters were approximately 12, 18, and 24 months, respectively, in the past, thus defining three retention intervals. At test session 2, these three intervals were increased by six months, and then again increased by six months at test session 3. In all, there were five different retention intervals, at 12, 18, 24, 30, and

36 months. The analyses below proceed first to examine differences across information types for intervals within each test session. This examination is followed by an analysis between sessions, and finally with an analysis using the five retention intervals.

For the analysis within each of the three test sessions (between semesters), there was a significant main effect for information type. The following *F* values were obtained: for the initial testing, $F(3,135) = 7.62$, $MSe = .51$; for the second testing, $F(3,108) = 10.31$, $MSe = .73$; and finally, for the third session, $F(3, 75) = 9.13$, $MSe = .62$. (Note: Analyses for semester and test session were done on proportions and thus mean square error terms are appropriate for proportions. Results, however, are reported in terms of percentages.) The critical finding was that for all three test sessions, recall performance for where information was significantly better than any of the other three types of information.

For the 48 subjects in the initial test session, mean percentages correct were 81.27%, 62.50%, 59.19%, and 60.77% for where, what, who, and when, respectively. Statistically, where performance was better in all cases [$F(1,45 = 23.93$, $MSe = 1.69$] for where/what, [$F(1,45) = 14.72$, $MSe = 2.34$] for where/who, and [$F(1,45) = 13.47$, $MSe = 2.02$] for where/when. For the initial testing, there was no main effect for semester. In this case, recall was tested either 12, 18, or 24 months following the

semester being questioned. Mean percentages correct were 57.14% for the Fall semester 1986, 66.42% for the Spring 1986, and 74.23% for Fall 1985. While there was no main effect for semester, it did appear that subjects did better on the Fall 1985 semester, which for most subjects was their first semester at the University. Statistically, performance on the Fall 1985 semester was not better than Spring 1986, but was better than Fall 1986 [$F(1,45) = 4.93$, MSe = .94]. The variable semester did not interact significantly with information type.

For the 39 subjects in the second test session, mean percentages correct were 72.90%, 45.46%, 43.56%, and 49.51% for where, what, who, and when, respectively. As in the first test, performance for where was significantly better than performance on any other information type [$F(1,36) = 34.67$, MSe = 2.91] for where/what; [$F(1,36) = 28.99$, MSe = 3.62,] for where/who; and [$F(1,36) = 8.47$, MSe = 1.72] for where/when. There was clearly no main effect of semester. In this case, recall was tested either 18, 24, or 30 months following the semester being questioned. Mean percentages correct were 53.96%, $n = 14$, for the Fall 1986 semester, 53.47%, $n = 9$, for Spring 1986, and 51.55%, $n = 16$ for Fall 1985. Unlike in the initial testing, there were no significant differences found among these semesters. In addition, semester did not interact significantly with information type.

Finally, for the 28 subjects in the third test session, mean

percentages correct were 69.39%, 42.57%, 36.36%, and 46.71% for where, what, who, and when, respectively. Again, performance for where information was significantly better than any other information type [$F(1,25) = 35.61$, $MSe = 2.15$] for where/what, [$F(1,25) = 22.46$, $MSe = 3.20$] for where/who, and [$F(1,25) = 11.61$, MSe 1.70] for where/when. Table 1 provides the mean percentages correct by information type for all three test sessions.

Table 1

Mean Percentage Correct Across Information Types by Test Session for Experiment 1

	Information type				
Session	what	who	where	when	
One (n = 48)	62.50	59.19	81.27	60.77	(65.93)
Two (n = 39)	45.46	43.56	72.90	49.51	(52.86)
Three (n = 28)	42.57	36.36	69.39	46.71	(48.76)

Note. Row means are provided in parentheses.

As in the second test session, there was no main effect for semester in this session. In this case, recall was either 24, 30, or 36 months following the semester being tested. Mean percentages were 47.94%, n = 8, for the Fall 1986 semester, 48.14%, n = 11, for Spring 1986, and 50.25%, n = 9, for Fall 1985. Though no significant differences were found among semester percentages, there was a Semester X Information Type interaction [$F(6,75) = 2.92$, $MSe = .20$]. Where information showed a significant effect for semester [$F(2,27) = 3.98$, $MSe = .48$]. Performance, in this case, improved with the shorter retention interval. For the other information types, no semester effect was found and there was a tendency for the best performance to be in the Fall 1985 semester (the longest interval). This interaction is readily apparent in Figure 1.

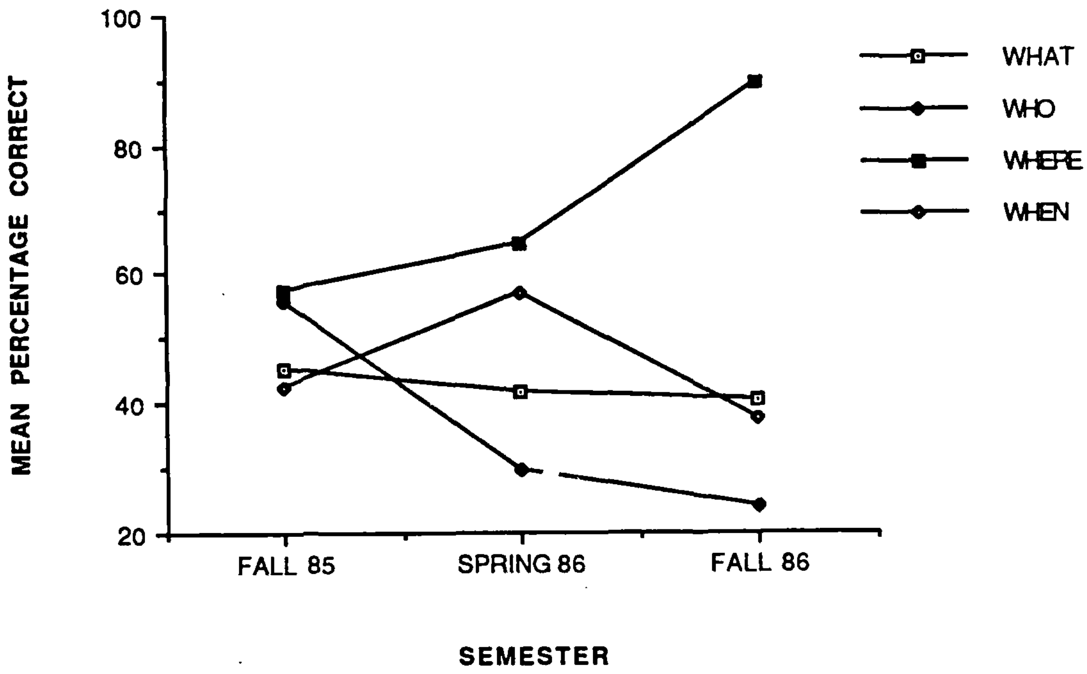


Figure 1. Semester X Information Type interaction for test session 3 in Experiment 1.

An additional analysis was performed which examined data across the three test sessions. In this case a 3 X 4 within-subjects ANOVA was conducted with test session (first, second, and third testing) and information type (what, who, where, and when) as the two factors. Twenty-six subjects were used in this analysis. (Note: This is the number of subjects who completed all three testings). Of interest here was whether there was a decrease in performance across testings. On average, the first test session used a retention interval of 18 months, the second session, 24 months, and the third session, 30 months. Overall percentages correct were 68.27%, 53.41%, and 48.26% for test sessions 1, 2, and 3, respectively. There was a main effect for test session [$F(2,50) = 8.83$, $MSe = 1.12$]. Single degree of freedom tests revealed that subjects performed significantly better during test session 1, the shorter retention interval (18 months) than for either of the other intervals [$F(1,25) = 8.50$, $MSe = 9.18$] for test 1 versus test 2 (24 months), and [$F(1,25) = 12.81$, $MSe = 16.66$] for test 1 versus test 3 (30 months). No significant difference was found between the second and third testing sessions. These results suggest that forgetting was greatest between the first and second year, but decreased somewhat after the second year.

In the analysis across test sessions, there was also a main effect for information type [$F(3,75) = 14.76$, $MSe = 1.35$]. Mean percentages correct for where, what, who, and when were 76.32%,

51.68%, 48.68%, 49.67%, respectively. Statistically, performance for where information was better than that for any of the other types of information [$F(1,25) = 50.27$, $MSe = 14.21$] for where/what, [$F(1,25) = 28.87$, $MSe = 17.56$] for where/who, and [$F(1,25) = 20.84$, $MSe = 16.62$] for where/when. No significant Test Session X Information Type interaction was found. There was a marginal interaction, however, for what and where information from test session 1 to test session 2 [$F(1,25) = 3.77$, $MSe = .46$, $p < .06$]. As can be seen in Figure 2, the decrease in performance from test session 1 to session 2 was least for where information.

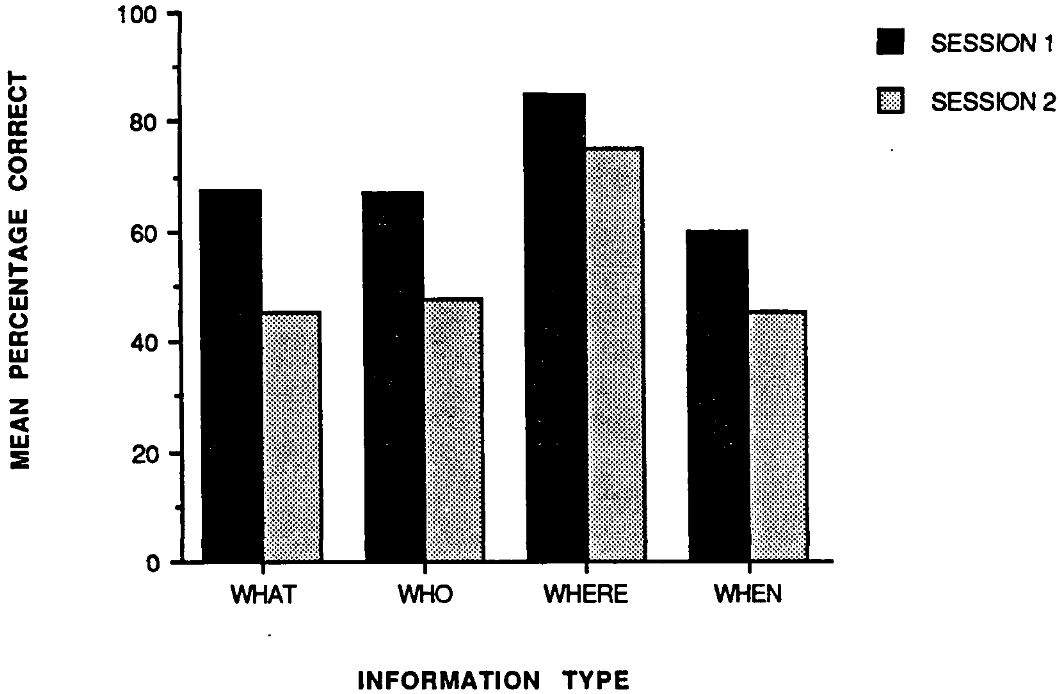


Figure 2. Test Session X Information Type interaction for what versus where information in Experiment 1.

This difference in performance was statistically significant only for where versus what information. No other interactions were were found to be significant.

In sum, the analysis of information type differences over time has revealed a consistent performance advantage for where information over what, who, and when information recall. This advantage was found within and across test sessions. Further, there was some evidence to suggest that where information was retained better over time than the other types of information. This was particularly true in the analysis across test sessions where it was found that performance dropped significantly from test session 1 to session 2 for all information types except where information.

Experience Ratings

In a natural learning study such as this, it is important to determine to what extent degree of learning contributed to results. Specifically, we wanted to know if the where advantage was attributable to more experience with campus locations over class times, professors, or courses. Subject experience ratings were used to examine this factor. As already discussed, subjects were asked to rate (on a 5-point scale from 0 to 4 or more) their experience with the what, who, where and when aspects of each course being tested. More specifically, subjects were asked, since the end of the test semester, how many times they took a

course in the same subject area, had the same instructor, had a course in the same building, and had a course at the same time. Ratings were summed across the three courses tested, making the range of possible scores from 0 to 12. A one-way within-subjects ANOVA was performed on these rating data to examine whether the where advantage in recall could be attributed to greater experience with where information over the other types of information.

In the first test session, the main effect of information type for experience was found to be significant [$F(3,138) = 43.06$, $MSe = 146.35$, , (Note: $n = 47$, one subject did not complete all ratings)]. Mean ratings were 3.64, 0.58, 4.47, and 4.02 for what, who, where, and when experiences, respectively. Experience with where information was significantly greater than that for what information [$F(1,46) = 5.18$, $MSe = 32.16$] and who information [$F(1,46) = 135.09$, $MSe = 710.66$]. There was no significant difference between where and when experience ratings. Experience with who information was clearly the anomoly in these data. Who ratings were significantly below what [$F(1,46) = 79.62$, $MSe = 440.46$], when [$F(1,46) = 85.35$, $MSe = 556.73$] and where, as noted above. This should not be surprising, however, since having the same professor for more than one class is an unusual experience at most universities. Additional analyses were performed examining the degree of correlation between experience ratings for a particular information type and

performance on that type of information. Correlation coefficients between experience and performance were .20, .01, .02, and .02 [F(1,44)] for what, who, where, and when, respectively. Only the correlation between what experience and what performance was significantly greater than zero [F(1,44) = 10.78, MSe = .53], though even this correlation was small.

Similar results were found in the second testing. The main effect for experience was significant [F(3,105) = 38.51, MSe = 176.00, (Note: n = 36, three subjects did not complete all ratings)]. Mean ratings were 4.11, 0.38, 4.86, and 5.17 for what, who, where, and when experiences, respectively. Experience with where information was significantly greater than only that for who information [F(1,35) = 101.98, MSe 722.53] and not different from experiences with what or when information. Experience with who information again was clearly less than that for what [F(1,35) = 83.84, MSe = 500.86], for when [F(1,35) = 96.20, 824.84], and for where, as already cited. Correlations between experience and performance in the second test were .10, .04, .09, and .01 [F(1,32)] for what, who, where, and when, respectively. All were nonsignificant.

Finally, results from the third testing were consistent with the first two sessions. The main effect for experience was significant [F(3,78) = 43.93, MSe = 192.79, (Note: n = 27, one subject did not complete all ratings)]. Mean experience ratings were 5.37, 0.85, 5.75, and 7.01 for what, who, where, and when,

respectively. Experience with where information was significantly greater than only that for who information [$F(1,26) = 92.11$, $MSe = 638.31$]. Experience with who information was again much less than that for what [$F(1,26) = 82.81$, MSe 542.98], when [$F(1,26) = 138.90$, $MSe = 1012.74$], and where information, as noted above. An interesting result in this analysis was the high ratings for when experiences. Though subjects tended to give high ratings for when experiences in earlier sessions, in this session, two semesters after starting the experiment, the mean when experience rating was significantly greater than who, as already noted, greater than what [$F(1,26) = 5.45$, MSe 72.62], and marginally better than where [$F(1,26) = 3.78$, MSe 43.02, $p < .06$]. This suggests that, despite the large number of possible class times (approximately 21), some times may be more likely to be experienced. Correlations between experience ratings and performance were .11, .03, .01, and .004 [$F(1,25)$] for what, who, where, and when, respectively. All were nonsignificant. Table 2 presents the mean experience ratings by information type for all three test sessions.

Table 2

Mean Experience Ratings Across Information Types by Test Session for Experiment 1

	Information type				
Session	what	who	where	when	
One (n = 47)	3.64	0.58	4.47	4.02	(3.18)
Two (n = 36)	4.11	0.38	4.86	5.17	(3.63)
Three (n = 27)	5.37	0.85	5.75	7.01	(4.75)

Note. Row means are provided in parentheses.

In sum, the analysis of experience ratings indicates that subjects did not receive a disproportionate amount of experience with where information over what or when information. Thus, the where advantage in test performance is not likely attributable to greater experience with where courses are held. In addition, the clear lack of repeated experiences with who information and the tendency for slightly more experiences with the same class times should be noted, though these differences are not readily reflected in performance scores.

Before turning to the next analysis, it is important to note that for each test session subjects were asked to indicate how many times they had reviewed their schedule (of the tested semester) since the semester ended. Almost all subjects said they had not reviewed their schedule. This is revealed in the mean number of review times of .23, .24, and .43 for tests 1, 2, and 3, respectively.

Enjoyment Ratings

An analysis was also conducted using the ratings subjects gave to how much they enjoyed a course, its professor, the building where it was held, and finally, the time it was held. These four questions were designed to coincide with the what, who, where, and when aspects of each course. As with experience ratings, enjoyment ratings were analyzed using a single factor, within-subjects ANOVA with the variable enjoyment having four

levels, corresponding to the four aspect questions above. Rating scores (on a scale from 1 to 7) were summed across the three courses being tested so that a minimum score was 4, indicating very little enjoyment, and the maximum was 28, indicating the highest degree of enjoyment. The interest in this analysis was whether the where advantage might be attributable to how much a subject enjoyed one aspect of a course over another.

In the first test session, the main effect of information type for enjoyment was not significant. Mean enjoyment ratings were 13.85, 13.68, 13.38, and 13.30 for the what, who, where, and when aspects of a course, respectively (Note: $n = 47$, one subject did not complete all the ratings). The where enjoyment rating was not significantly different from any of the others. Correlation coefficients were also produced comparing recall performance for each information type with the corresponding enjoyment rating, for example the relationship between recall performance for what information and the rating of course enjoyment. Performance-enjoyment correlations were .02, .03, .001, and .05 [$F(1,44)$] for the what, who, where, and when aspects of course. None of the correlations were significantly greater than zero. Finally, in addition to rating how well they enjoyed each aspect of a course, subjects rated how well they enjoyed the entire test semester. The mean semester enjoyment rating was 4.73 (out of 7). The correlation between overall recall performance and the semester enjoyment rating (.05) was

not significant. Correlations between recall performance by information type and semester enjoyment rating were also nonsignificant. They were .07, .02, .02, and .01 [$F(1,44)$] for what, who, where, and when aspects, respectively.

For the second session, the main effect for course enjoyment ratings was also not significant. The mean enjoyment ratings were 13.92, 13.33, 13.46, and 12.69 for the what, who, where, and when aspects of a course ($n = 39$). Also as in test one, enjoyment ratings for where information were not different from ratings for the other types of course information. Correlations between recall performance and enjoyment ratings for each type of information were not significant as well. They were .06, .04, .04, and .08 [$F(1,32)$] for the what, who, where, and when aspects, respectively. Interestingly though, rating of semester enjoyment, in this test session, was significantly related to overall recall performance [$F(1,36) = 11.71$, $MSe = .40$], with a correlation of .25. The mean semester enjoyment rating was 4.46 out of 7, $n = 37$ (two subjects did not complete this rating). Correlations by information type and semester enjoyment were marginally significant or greater [$F(1,32) = 3.76$, $MSe = .21$, $p < .06$] for what information, [$F(1,32) = 4.86$, $MSe = .52$] for who, [$F(1,32) = 5.94$, $MSe = .40$] for where, and [$F(1,32) = 5.60$, $MSe = .67$] for when. Correlations were .11, .13, .16, and .14 [$F(1,32)$] for what, who, where, and when, respectively.

Consistent with the other two sessions, the main effect of

information for enjoyment was not significant in the third test session. Mean enjoyment ratings were 13.86, 12.93, 13.89, and 13.61 for the what, who, where, and when aspects of a course ($n =$ 28). Also consistent with earlier tests was the lack of any significant differences between ratings for where information and the other types of information. Similarly, correlations between recall performance and enjoyment ratings for each type of information were not significant. Correlation coefficients were .01, .0001, .02, and .003 [$F(1,25)$] for what, who, where, and when, respectively. Unlike in test session 2, but consistent with session 1, the correlation between overall recall performance and the semester enjoyment rating (.009), was not significant. The mean semester enjoyment rating was 4.75. Likewise, none of the four types of information were correlated with the semester enjoyment rating. Correlations were .02, .001, .008, and .0002 [$F(1,25)$] for what, who, where, and when, respectively. Table 3 presents mean enjoyment ratings by information type for all three test sessions.

Table 3

Mean Enjoyment Ratings Across Information Types by Test Session for Experiment 1

	Information type				
Session	what	who	where	when	
One (n = 47)	13.85	13.68	13.38	13.30	(13.55)
Two (n = 39)	13.92	13.33	13.46	12.69	(13.35)
Three (n = 28)	13.86	12.93	13.89	13.61	(13.57)

Note. Row means are provided in parentheses. Ratings could range from 4 to 28.

In sum, results of the analyses of enjoyment ratings suggest that the advantage found for where information in recall cannot be attributable to differences in ratings of how much subjects enjoyed the four aspects of their courses. There was some indication that rated enjoyment of the entire semester was correlated with recall level, but in this case, significant correlations were found for overall performance.

Gender

An analysis of information type by gender performance was conducted assuming results might help in understanding the locus of the where advantage in performance. In the first test session using a full 48 subjects, there were 19 female and 29 male subjects. The analysis revealed a marginally significant effect for gender, with female subjects scoring better than males [$F(1,46) = 3.42$, MSe .66, $p < .07$]. Mean percentage correct for female subjects was 73.16% and 61.20% for male subjects. There was no significant Gender X Information Type interaction. The general trend was for female subjects to do better across all four types of information. However, differences between female and male performance were statistically significant only for where and when information [$F(1,46) = 4.67$, MSe = .31] for where and [$F(1,46) = 4.02$, MSe = .46] for when. In both cases female subjects scored higher than male subjects. Mean percentages for where were 91.26% for female subjects and 74.72% for male

subjects. Percentages for when information were 72.84% for female and 52.86% for male subjects.

In the second test session, with 39 subjects participating, there were 16 female and 23 male subjects. In this case, the analysis yielded a significant main effect for gender with female subjects doing better than male subjects [$F(1,37) = 4.22$, $MSe = .16$]. Mean percentages correct were 60.81% and 47.33%, for female and male subjects, respectively. As in test one, there was no significant interaction between gender and information type. The data suggest better performance for female subjects across all four information types. These differences, however, reached statistical significance for what information only, with female subjects performing at 56.25% and male subjects at 37.96% [$F(1,37) = 5.80$, $MSe = .32$].

Finally, in the third test session, with 28 subjects participating, there were 10 female and 18 male subjects. Again, there was a significant main effect for gender [$F(1,26) = 4.18$, $MSe = .76$]. Female subjects, with a mean percentage of 59.83%, performed better than male subjects with a mean percentage of 42.61%. There was also no significant interaction between gender and information type. As before, female subjects tended to do better on all four types of information. In the case of what and when information, differences across gender reached statistical significance [$F(1,26) = 5.76$, $MSe = .31$] for what, and [$F(1,26) = 4.31$, $MSe = .62$] for when. Table 4 provides mean percentages by

gender and information type for all three test sessions.

Table 4

Mean Percentage Correct Across Information Types by Gender for Experiment 1

Session	Information type: what	who	where	when	
One					
male (n = 29)	62.93	54.28	74.72	52.86	(61.20)
female (n = 19)	61.84	66.68	91.26	72.84	(73.16)
Two					
male (n = 23)	37.92	39.48	69.22	42.65	(47.33)
female (n = 16)	56.25	49.44	78.19	59.38	(60.81)
Three					
male (n = 18)	34.72	34.33	65.78	35.61	(42.61)
female (n = 10)	56.70	40.00	75.90	66.70	(59.83)

Note. Row means are provided in parentheses.

In sum, for all three test sessions female subjects did better than male subjects. It does not appear, however, that this gender effect is related to information type. Across tests, female subjects performed significantly better than male subjects for what, when, and where information. Though the superior performance of female subjects did not reach statistical significance for who information, a trend in that direction was evident.

Information Types as Cues

In addition to looking at differences in how well each type of information was remembered, an analysis was performed examining each type of information as a cue for recall. Two questions were being addressed in this analysis. First, was where information distinct from the other types of information in terms of cuing effectiveness. Secondly, was the superior performance of where information in recall related to any particular recall cue. In this analysis, individual subjects were combined into groups of four so that in each group of subjects, all twelve cue-information combinations were found, three combinations per subject. Thus, in each test session, 12 such grouped observations were constructed. Analyses were conducted only for performance after one cue was given (the primary cue). Preliminary analysis of performance after two cues (the primary and secondary cues) revealed no significant effects.

Results from the first session yielded a nonsignificant main effect for cue type [$F(3,33) = 2.10$, $MSe = .05$, $p < .12$]. Mean percentages correct for each cue were 72.72%, 71.36%, 61.08% and 60.22% for what, who, where, and when cues, respectively. Comparisons made between the where cue and each of the other cues were also not significant. To address the second issue in this analysis, whether the where advantage was unique to a particular cue type, separate analyses were performed for each cue type. Since a particular information type was never matched with its own cue, a full interaction model using all four cue and information types could not be tested. Thus, individual one-way ANOVAs were conducted with information type (3 levels only) as the single within-subjects factor.

For the what cue, there was a main effect for information type [$F(2,22) = 9.78$, $MSe = .32$]. Mean percentages were 91.67%, 62.50%, and 64.00% for where, who, and when information, respectively. Where information recall was significantly better than both who [$F(1,11) = 16.56$, $MSe = .102$] and when information [$F(1,11) = 15.88$, $MSe = .92$]. For the who cue, the main effect for information type was not significant. Performance for where information (80.58%) appeared better than what (68.00%) and when information (65.55%), but was not significantly so. For the when cue, there was a main effect for information type [$F(2,22) = 5.32$, $MSe = .13$]. Mean percentages were 72.33%, 52.75%, and 55.58% for where, what, and who information, respectively. In

this case, where information recall was significantly better than what [$F(1,11) = 15.55$, $MSe = .46$] and who information [$F(1,11) = 4.71$, $MSe = .34$]. These results suggest that the where advantage in recall cannot be easily attributed to the effect of one particular cue type over another. Table 5 provides mean recall percentages by cue type and information type for test session 1.

Table 5

Mean Percentage Correct Across Information Types by Cue Type in Test Session 1 of Experiment 1

	Information type				
Cue	what	who	where	when	
what	_____	62.50	91.67	64.00	(72.22)
who	68.00	_____	80.58	65.50	(71.36)
where	66.58	66.67	_____	50.00	(61.08)
when	52.75	55.58	72.33	_____	(60.22)

Note. Row means are provided in parentheses.

A cue type was never matched with the same information type.

In the second testing, 12 combined observations were constructed, however, 9 observations had data from only 3, not 4 subjects. Results yielded a marginally significant overall effect of cue type [$F(3,33) = 2.80$, $MSe = .11$, $p < .06$]. Mean percentages correct for each cue were 62.28%, 58.58%, 44.19%, and 44.36% for what, who, where, and when cues, respectively. Comparisons made between the where cue and the others showed the what to be significantly better than where [$F(1,11) = 9.30$, $MSe = .39$], but a lack of significant difference between the where cue and the who and when cues. Individual ANOVAs examining recall differences across information type revealed results similar to those found in test session 1. For the what cue, there was a main effect for information type [$F(2,22) = 7.06$, $MSe = .61$]. Mean percentages were 86.75%, 42.42%, and 57.67% for where, who, and when, respectively. Where information recall was found to be significantly better than who [$F(1,11) = 20.75$, $MSe = 2.36$] and when information [$F(1,11) = 5.60$, $MSe = 1.02$]. Similarly, for the who cue, a main effect for information type was found. Mean percentages were 82.00%, 53.50%, and 40.25% for where, what, and when, respectively. Again, where information was better recalled than what [$F(1,11) = 18.64$, $MSe = .98$] and when information [$F(1,11) = 12.60$, $MSe = 2.09$]. The main effect for information type in the when cue analysis was not significant. Mean percentages were 55.50%, 36.00%, and 41.58% for where, what, and who information, respectively. Where information, however, was

found to be marginally better recalled than what information [$F(1,11) = 3.87$, $MSe = .46$, $p < .07$]. Table 6 provides mean percentages correct by cue type and information type for test session 2.

Table 6

Mean Percentage Correct Across Information Types by Cue Type in Test Session 2 of Experiment 1

	Information type				
Cue	what	who	where	when	
what	_____	42.42	86.75	57.67	(62.28)
who	53.50	_____	82.00	36.58	(58.58)
where	47.92	38.92	_____	45.75	(44.19)
when	36.00	41.58	55.50	_____	(44.36)

Note. Row means are provided in parentheses.

A cue type was never matched with the same information type.

Finally, in the third test session only 8 of the combined observations could be constructed, with 4 of these having data from only 3 subjects. Results yielded a significant main effect for cue type [$F(2,21) = 7.01$, MSe = .24]. Mean percentages were 60.75%, 63.58%, 29.08%, and 35.46% for what, who, where, and when cues respectively. Comparisons made between the where cue and both what and who cues were significant [$F(1,7) = 24.90$, MSe = .80] for what and [$F(1,7) = 11.18$, MSe = .95] for who. Both what and who cues produced better performance than did the where cue. The where and when cues did not differ significantly. Individual ANOVAs by cue type revealed a significant main effect for information type only for the who cue [$F(2,14) = 6.38$, MSe = .15]. Mean percentages were 79.25%, 55.13%, and 56.38% for where, what, and when, respectively. No main effect for information type was found for the what and when cues. Single degree of freedom comparisons of where information with the other information types were also not significant. The trend for these cues, however, was for where information to show the best recall performance. Table 7 provides mean percentages correct for all cues by information type for test session 3.

Table 7

Mean Percentage Correct Across Information Types by Cue Type in Test Session 3 of Experiment 1

Cue	Information type: what	who	where	when	
what	____	51.12	83.25	47.88	(60.75)
who	55.13	____	79.25	56.38	(63.58)
where	35.38	29.13	____	22.75	(29.08)
when	33.38	32.38	40.63	____	(35.46)

Note. Row means are provided in parentheses.

A cue type was never matched with the same information type.

In sum, the analysis of retrieval cues suggests that differences do exist among information types in their usefulness as cues during recall. The where cue along with the when cue appear to be less effective than the what and who cues. In addition, comparisons within cue types suggest that the advantage for where information in recall is not a phenomenon peculiar to one cue type. Instead, the where advantage is seen for all three cues, the what, who, and when cues.

Discussion

All three test results in this experiment have demonstrated a clear advantage in recall of where information over recall of what, who, or when information. Analysis of recall performance over time has suggested that where information retains its advantage for several years, even as the other types of information show substantial loss. This advantage cannot be attributed simply to subjects having more experience with where a course was held than with courses in the same subject area, with the same professor, or at the same time. Likewise, where information recall was not superior simply because subjects enjoyed the location of a course better than the course material, the professor, or the course time. In the second test session, higher ratings of semester enjoyment were correlated with better recall performance, but for overall levels, not just where. Likewise, both White (1982) and Wagenaar (1986) found t

ratings of event pleasantness were correlated with event recall performance. Results of the experiment did show better performance for female subjects than for male subjects, but female subjects were better for where, what, and when information. No unique advantage for where information was found. In like manner, Robinson (1976) in a probe word study, found that female subjects recalled more recent events versus farther in the past, than male subjects. In addition, the analysis of cue effectiveness revealed that where information was a relatively poor cue, but not uniquely so. Both where and when information were equally poor cues and were less effective than what and who information as cues for recall. Finally, the cue data suggest the where advantage cannot be attributed simply to one type of recall cue being particularly effective at soliciting where information. Recall of where information was superior regardless of which cue was used. The question then, still remains. What is the locus of the where advantage?

It can be argued that the advantage found for where information is due to some unique characteristic of the where recall task. In the questionnaire, recalling what, who, and when information involved reading a question and then filling a blank space with words or numbers. More specifically, completing a what question required producing a relatively meaningful, one to four word course title. A who response involved producing a one-word name having considerable name-face association. A when

response required recall of a time. These three recall tasks can be contrasted with the requirements of the where task. The where task involved no verbal response, only a simple mark next to the building where the class was held. Each mark was placed on a campus map, void of verbal labels, i.e., street and building names, etc. The where recall task then, was unique in two important respects: it provided subjects with a map of the campus (more than a blank line), and required the use of spatial knowledge.

Providing the campus map may have helped recall performance in at least two ways. It may have helped recall in that subjects could have used the features depicted on the map, i.e., buildings, roads, trees, as additional cues for recall. This is referred to as the multiple cues hypothesis. Having these additional cues could have given where recall an advantage. It could be said that subjects simply examined the map for familiar features in hope that the correct location of a course would come to mind. The problem with this simple account of a map cuing effect is that subjects are likely to find many familiar features on the map. Unless these features were somehow associated with the particular course in question, they would be of little help.

Another advantage to having the map was that it could have made the where task a recognition memory task rather than a recall task (the recognition hypothesis). With the map, subjects were provided with all the possible responses to a where

question. The map presented subjects with a visual display of all the buildings on campus. Of the over 100 locations, one of them was the correct one of the course in question. In actuality, there were fewer alternatives. Given some familiarity with the campus, a number of buildings were not likely locations, i.e., dormitories, administration buildings. In a sense, this task can be viewed as a forced choice recognition task where subjects simply recognized locations from among the likely possibilities. In contrast, in the what, who, and when recall tasks subjects had to recall explicitly the information. Consequently, the where task, involving recognition memory, yielded better recall. As a counter argument, it could be claimed that the when task was similar to the where task in this respect. For the when task, the possible class times (approximately 21 of them) were available to subjects much like the buildings on the map. However in this case, the alternatives were not presented visually, but were part of a subject's general knowledge about class schedules. It is common student knowledge, especially for juniors and seniors, as used in this experiment, that classes start on the hour or half hour, depending on the day of the week. Consequently, as with a where response, subjects could select when responses from among a relatively finite set of alternatives. Despite this similarity in tasks, however, where information was still better recalled than when information. Perhaps this similarity between where and when task (having a

relatively finite set of alternatives) best explains cuing results in which what and who cues, being more unique, served as better cues.

The second defining characteristic of the where task was its spatial nature. Subjects had to locate each course on a fairly detailed two-dimensional map of the campus. It can be argued that spatial knowledge could have contributed to the where information advantage in three ways.

First, spatial knowledge may have contributed to the where advantage in the form of procedural information, or in knowing how to proceed through the campus space (the spatial procedures hypothesis). In getting from class to class during a semester, subjects are likely to learn specific routes through the campus and come to associate these routes with specific buildings and other campus features. Following Kolers and Roediger (1984), it might be said that subjects, with the aid of the map, reinstate previously learned procedures at test. Performance for the what, who, and when tasks was not as good because these tasks involved less reinstatement of learned procedures. Assuming that knowledge can be both declarative and procedural, it could also be said that over the course of a semester, subjects lose their declarative knowledge of their course schedule (the what, who, where (building name), and when), but develop and maintain their procedural knowledge of their schedule (getting to and from their classes). In this case, the where task taps into the

proceduralized knowledge, thus showing an advantage at recall (see Anderson, 1982; Cohen, 1984). Secondly, the spatial knowledge advantage may have been due to the use of both visual and verbal information, as opposed to the largely verbally coded information found in the other tasks (Paivio, 1971, 1986). It could be argued that because the map provided a visual representation of the campus, it facilitated visual imaging of the campus by the subjects. Thus, subjects may have formed visual images of the campus, internally sorting through them until recognizing the features, i.e., buildings and other objects, associated with the course in question (Kosslyn, 1976, 1987). The where advantage would then be due to having available at recall visual information, cued via the map and any concommitant verbal information. The what and when tasks likely had minimal visual components and the who task may have involved some visual imaging but did not have the benefit of the map as cue. Thirdly, to locate properly the correct buildings on the campus map, subjects may have used their knowledge of spatial relationships between buildings and other campus landmarks, along with knowledge of relative distances among these objects. Following the spatial retention hypothesis, it could be argued that spatial relations information is simply retained better in long-term memory than the largely verbal information required in the other tasks. Unique processing of spatial information has been suggested by Healy (1978, 1982).

In summary, the consistent advantage for where information found in this experiment could be accounted for in five ways. According to the multiple cues and recognition hypotheses, the map, by simply providing additional cues or by making the task a multiple choice recognition task, produced a where information advantage. Three additional hypotheses (the procedural, visual coding, and spatial retention hypotheses) were also proposed, each emphasizing the role of spatial information in the where advantage.

To examine these possibilities, a second experiment was planned. This experiment was designed so that subjects would learn a semester schedule in a laboratory setting. By moving into the laboratory several hypotheses could be evaluated. First, performance could be examined when learning did not actually involve proceeding to classes. Subjects in this experiment learn the what, who, where, and when of an actual schedule, but never implement their schedule. If subjects continue to do better on where information recall, a proceduralization of knowledge account such as proposed by Anderson (1982) would need modification. Secondly, the spatial relations and visual imaging properties of the map could be examined more closely by constructing two recall tasks for where information. One task simply used the campus map as in Experiment 1. The other task was a verbal task which required subjects to name the buildings where classes were held. If the

where advantage was found for both the spatial location and location naming tasks, a spatial knowledge account of the advantage is not likely.

The second experiment was also designed to determine if the where advantage was due to some natural learning factor, beyond the explanations already considered. For example, one factor considered was the effect of spacing learning across a full semester. It is possible that repeated practice over many weeks had a differential effect on the four types of information. Subjects therefore, received either massed training, all in one day, or had training spaced across three weeks. Finding a where advantage for spaced trained subjects and not for massed trained subjects would suggest that extended training might be a factor in the where advantage. A second training factor examined was the degree of practice each type of information received. In studying long-term retention of Spanish learned in school, Bahrick (1984) found that degree of original learning was a major predictor of memory performance. In the laboratory all four types of information would receive equal study. In the real-world, learning where and when a class is held may receive greater emphasis or practice over a semester than would learning what a course is called and who is teaching the course. It is possible in the case of where information, that knowledge necessary for getting to classes received greater practice than knowledge of building names. Similarly, subjects may have

practiced simply the sequencial order of their classes during the school week more than knowledge of the actual start times. Thus, an additional manipulation in this experiment was to give one group of subjects extra practice in recall of building names and class order, and another group, extra practice recalling building locations and class start times. If spatial knowledge is implicated in the where advantage, subjects practicing building locations more than names would likely show a stronger where advantage at recall.

In sum, by controlling in the laboratory what kinds of information are studied, and to what extent they are studied, the locus of the where advantage can be more readily identified. The full details of this experiment are provided in the next chapter.

CHAPTER III

EXPERIMENT 2

Method

Subjects

As in Experiment 1, 48 University of Colorado undergraduates enrolled in introductory psychology participated in this study for course credit. Also as before, subjects were restricted to only those students who had attended the University for at least two years, essentially, only juniors and seniors.

Materials

Subjects were asked to learn a fictitious class schedule. These schedules were constructed from the actual schedules of subjects used in Experiment 1. The fictitious schedules contained either 4 or 5 courses (24 of each). The three courses used in the questionnaire in Experiment 1 were always used in the fictitious schedules. The remaining one or two courses were constructed using the same student's schedule when enough courses were available, or by taking unused courses from another student's schedule. Missing information was completed using a university listing of courses. In sum, each fictitious schedule was made up of courses, buildings, times, etc., found in real

schedules.

Class schedules were in two parts. The class listing looked much like a standard university class schedule (see Appendix C). Classes were described, one per line, using eight column headings; department, course number, course title, instructor, class start and finish time, days of the week, building name, and room number. The department, days, and building names were always abbreviated, and at times, so was the course title, just as in actual schedules. The second part of the schedule was the class map. The map was a two-dimensional, black and white drawing of the University campus (much like that shown in Appendix A). It included trees, shadowing (gave 3-D appearance), street names, and an index to the building names (each building was marked with a number which matched a directory of names and location coordinates). The buildings where classes were held were marked with blank yellow circles (numbers were whited out). The map also showed the relative location of the Rocky Mountains to the campus and the direction of North.

During the training phase of the experiment, subjects were asked to recall their schedule information based on their task assignment, either the class listing or map task. The class listing task involved completing the missing information on a form much like the class schedule used during study (see Appendix D for sample form). The form, however, differed in that it had only six rather than eight headings: department, course number,

course title, instructor, class time (but no days of the week), and building name (but no room number). In addition, the class listing forms had four columns left blank; course title, instructor, class time, and building (what, who , when, and where information). The department and course number were provided and served as recall cues. Those subjects assigned to the map task used a slightly modified form (see Appendix E). On this form subjects marked the buildings on a map indicating where their classes were held. The maps in this task were the same as the maps used during study except the numbers on all buildings were omitted and no building name directory, nor street names were included. The map continued to show the location of the Rocky Mountains and the direction of North. In addition to completing the map (where information), a modified class listing was included on the map form. The five column headings on the map task form included department and course number (which were filled in), and course title, instructor, and class order (what, who, and when information), which required completion. The class order heading required listing of the order of classes during the school week, first, second, third, etc.

In addition to the new materials used in this experiment, the class schedule questionnaire from Experiment 1 was also used for testing recall. The same cue-then-question format was used wherein subjects were given one then two cues in recalling the what, who, where, and when information for each of three courses.

The questionnaire used in this experiment only differed in the questions asked following recall. Instead of questions about course and semester experiences, this version of the questionnaire focused on the differences and similarities subjects found between studying their fictitious schedule and studying their own schedule in real life. The specific questions can be found in Appendix F.

Design and Procedure

A 2 X 2 X 2 X 4 mixed factorial design was used combining the following between-subjects variables: (1) training type (spaced or massed), (2) orienting task (class listing or map), and (3) number of courses (4 or 5). Information type (what, who, where, or when) was the single within-subjects factor. All subjects received nine study-then-recall training trials spread across three sessions (spaced training) or grouped in one session (massed training). During week four of the experiment all subjects were administered the class schedule questionnaire followed by both the map and class listing recall tasks. Approximately five weeks following this first test session, the questionnaire and both the map and class listing tasks were again administered.

Subjects in all conditions were told the experiment was designed to examine how well students remember their university class schedules. They were informed that during the course of

the experiment they would repeatedly study and then recall a fictitious schedule. They were instructed to study their schedule as if they were learning their own schedule at the start of the semester.

In the training phase of the experiment, subjects were given nine study trials where they were asked to study both their class listing and class map. Each trial lasted five minutes. Subjects were told that each trial would be followed by a recall test. Precisely how recall would be tested was not specified. Subjects were simply instructed to study their schedule as if it was their own. The type of recall test was determined by which orienting task the subject was assigned to, either the class listing or map task. Those assigned to the class listing task were asked to fill in a blank class listing following eight of the nine study trials. Following one study trial, the fifth, these subjects were asked to complete the map task. Including one trial with the alternate task was done to discourage subjects from completely ignoring, during study, the information not being tested during recall. In the class listing task, for example, the subjects could neglect studying the map and class order information. For the subjects assigned to the map condition, the reverse was true. Subjects could neglect study of building names and class start times. Therefore, in the map task, a map form was used for recall on eight trials, with a class listing form used after the fifth study trial. Subjects were given a maximum

of five minutes to complete their recall, though few took this long after the first two study trials.

As for the type of training, subjects were randomly assigned to either the massed or spaced training condition. The 24 subjects receiving massed training were given all nine study/recall trials in a one and a half hour period. Brief breaks (approximately five minutes each) occurred after the third and sixth trials. Subjects were allowed to stand up, go to the restroom, etc., during breaks, but not discuss the experiment. These subjects were divided into three groups, with eight subjects receiving training either one, two, or three weeks prior to the test trial in week four. This was done to equate the average retention interval for massed training with that received by subjects in the spaced training condition. The 24 subjects in the spaced condition received their 9 study/recall trials, 3 per week, in 30 minute sessions, spaced over 3 weeks. Testing occurred on the fourth week. Both the massed and spaced groups had training trials separated into one-week intervals (precisely seven days). However, six subjects missed a scheduled training session and came one day late (five in spaced, one in massed) and two subjects had to reschedule one day earlier than their scheduled session (both in the spaced group).

In the testing phase of the experiment, both training groups followed an identical sequence. All subjects were first asked to complete the class schedule questionnaire. The questionnaires

followed the cued recall format described above. Three courses were randomly selected from each class schedule for testing. Following completion of the questionnaire, subjects were asked once again to complete both a blank class listing and a map, with the order of these two tasks being counterbalanced across subjects. With these tasks completed, subjects were thanked and dismissed. Test trials lasted approximately 30 minutes. All subjects were scheduled for testing precisely one, two, or three weeks after completing the last training trial, depending on their assigned training condition. However, five subjects missed their scheduled test time and were tested one day late (four in spaced training, one in massed). Approximately five weeks after testing, subjects were asked to participate in an unexpected retest session following the same procedures used in the first testing. Subjects were not informed of the precise nature of this last session, but were simply told that it would be similar to earlier sessions. Subjects had already met course requirements at time of retest and were thus paid five dollars for participating. In spite of upcoming final exams, 36 of the 48 original subjects volunteered to take part in retesting.

Results

The questionnaire, map, and class listing test, each retest, and the training type data were scored for percentage of correct responses using the same criteria as in Experiment 1. Briefly reviewing, scoring was quite strict, emphasizing consistency across information types. Responses received a score of 1 if they were perfectly accurate or very close. Partially correct responses received only half credit. Unique to this experiment was the requirement to provide building names (in the class listing test) and class order (in the map test). Full credit for building names were given only when the precise name provided in the schedule was used. Shortened names or misspellings received half credit. When subjects were required to provide the temporal order of classes during the school week, full credit was given only for the precise order. Half credit was given when the order was off by one, for example, a class listed as 4th was actually 5th.

Unless otherwise stated, a 2 X 2 X 2 X 4 analysis of variance (ANOVA) was performed for each test and retest and for the training data, with training type, orienting task, and number of courses as between-subjects factors and the four types of information as the single within-subjects factor. All statistical tests used a .05 level of confidence. For the questionnaire, map, and class listing data that follow, results

are first presented for the initial test and then followed with the retest results.

Analysis of Questionnaire Test Data

Results are summarized in Table 8 in terms of mean percentages correct as a function of information type, training, task, and number of courses.

Table 8

Mean Percentage Correct for Questionnaire Test Data Across Information Types as a Function of Training, Task, and Number of Courses

	Information type *				
	what	who	where	when	
Training *					
Massed	55.21	53.79	77.46	44.08	(57.64)
Spaced	85.75	89.25	82.67	85.08	(85.69)
Task *					
Map	64.58	62.50	77.46	58.67	(65.80)
Class Listing	76.38	80.54	82.67	70.50	(77.52)
Courses					
4	70.13	73.63	78.13	67.71	(72.40)
5	70.83	69.42	82.00	61.46	(70.93)
	(70.48)	(71.52)	(80.06)	(64.58)	(71.66)

Note. Number of subjects per cell was 24. Column and row means are provided in parentheses.

* Significant main effect at $p < .05$.

Training type (massed or spaced) proved to be highly significant [$F(1,40) = 21.83$, $MSe = 3.78$], with spaced training ($M = 85.69\%$) yielding better recall performance than massed training subjects ($M = 57.64\%$). (Note: Though tables and figures report percentages, all analyses in this experiment were done on proportions, thus, all mean square errors are appropriate for proportions). Training type did not interact significantly with orienting task or number of courses, but did with information type [$F(3,120) = 5.83$, $MSe = .30$], see Figure 3.

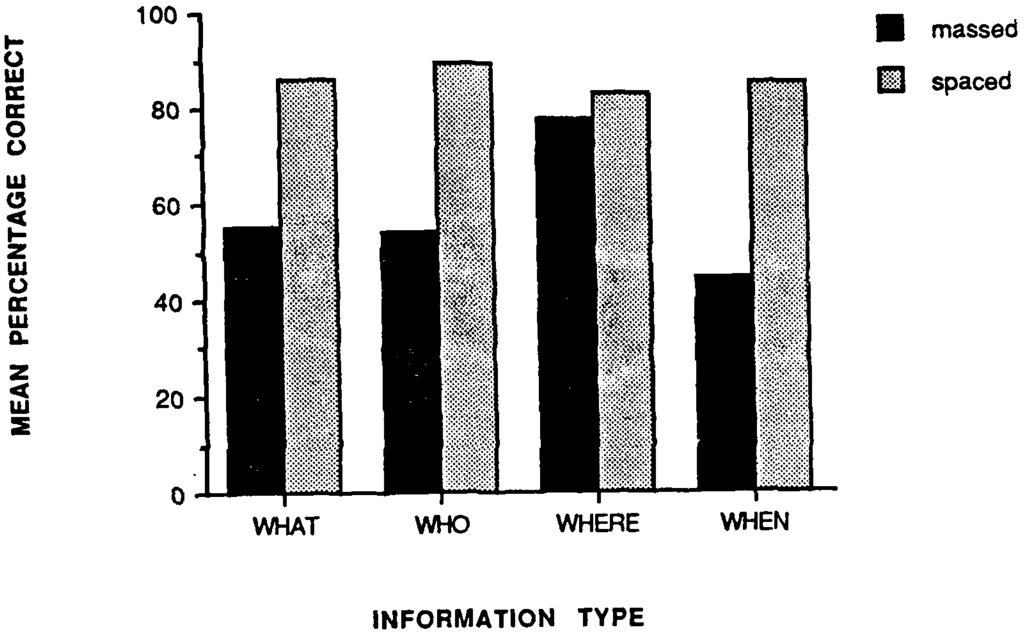


Figure 3. Information Type X Training Type interaction in the questionnaire test data, Experiment 2.

Separate 2 X 2 X 2 ANOVAs for each information type revealed significant training type differences for all but where information, with spaced training yielding the best performance, for what [$F(1,40) = 15.72$, $MSe = 1.12$], for who [$F(1,40) = 22.08$, $MSe = 1.51$], and for when [$F(1,40) = 19.75$, $MSe = 2.20$]. For where information spaced and massed subjects showed no significant difference in performance (Ms = 82.67% and 77.46%, respectively [$F(1,40) = .38$, $MSe = .032$, $p < .54$].

Orienting task also had a moderately significant effect on performance [$F(1,40) = 3.81$, $MSe = .66$], with subjects training under the class listing task performing better than those using the map completion task (Ms = 77.52% and 65.80% correct, respectively). This task type difference did not interact significantly with training or information type, but it did vary with the number of courses in a student's schedule [$F(1,40) = 5.28$, $MSe = .91$]. Differences in orienting task were evident only for subjects with five courses in their schedule and not for those with four courses. With five courses, performance was at 83.69% for the the class listing task and only 58.17% in the map task. With four courses, performance remained at the same level across tasks (Ms = 71.35% and 73.44% correct for the class listing and map tasks, respectively).

The main effect of number of courses in the schedule (4 versus 5) was not significant. The mean percentage correct for four courses was 72.40%, and 70.93% with five courses. Number of

courses did interact with task type as described above. No other interactions with the number of courses variable were significant.

Of primary interest in this study was performance differences among information types. The main effect of information type proved to be significant [F(3,120) = 3.80, MSe = .20]. Of more interest were the comparisons of where information with performance on the other three types of information. Consistent with Experiment 1, where information was found to have the best recall performance. The mean percentages correct for where, who, what and when information were 80.06%, 71.52%, 70.48%, and 64.58%, respectively. Performance for where information was significantly higher than performance on any of the other information types, for what [F(1,40) = 6.22, MSe = .44], for who [F(1,40) = 5.38, MSe = .35], and for when [F(1,40) = 6.90, MSe = 1.15]. As described earlier, there was an Information X Training Type interaction (see Figure 3). All but where information showed a training type effect. No other interactions with information type were significant.

Analysis of Questionnaire Retest Data

In the questionnaire retest, 36 subjects were tested again using the same three courses found in the initial testing. The same cue-question combinations were used, but in a different order. Despite retesting with the same courses, the mean overall

performance on the questionnaire retest was only 58.75%. This was significantly below the mean performance on the initial test, 72.10%, for the same 36 subjects [$F(1,28) = 15.33$, $MSe = .60$]. Table 9 provides mean percentages correct across information types by training, task, and number of courses.

Table 9

Mean Percentage Correct for Questionnaire Retest Data Across Information Types as a Function of Training, Task, and Number of Courses

	Information type *				
	what	who	where	when	
Training *					
Massed (n = 19)	56.58	51.79	55.74	33.79	(49.47)
Spaced (n = 17)	64.24	74.00	74.00	64.24	(69.12)
Task					
Map (n = 15)	70.60	59.47	77.27	42.80	(62.53)
Class Listing (n = 21)	52.76	64.29	55.14	52.00	(56.05)
Courses					
4 (n = 18)	51.39	66.67	62.50	51.44	(58.00)
5 (n = 18)	69.00	57.89	66.22	44.89	(59.50)
	(60.19)	(62.28)	(64.36)	(48.17)	(58.75)

Note. Column and row means are provided in parentheses.

* Significant main effect at $p < .05$.

As in the test data, training type proved to be a significant factor [$F(1,28) = 7.94$, $MSe = 1.35$]. Subjects undergoing spaced training ($M = 69.12\%$, $n = 17$) performed better than those given massed training ($M = 49.47\%$, $n = 19$). Training type did not interact significantly with task or number of courses. More importantly, there was no interaction of training with information type, as was found in the test data. This suggests that the test data Training X Information Type interaction was likely due to a ceiling effect for where information.

For orienting task, no main effect was found. Mean performance for the map task was at 62.53%, $n = 15$, and 56.05%, $n = 21$, for the class listing task. In spite of no main effect, orienting task did interact significantly with information type [$F(3,84) = 2.96$, $MSe = .20$]. This interaction is displayed in Figure 4.

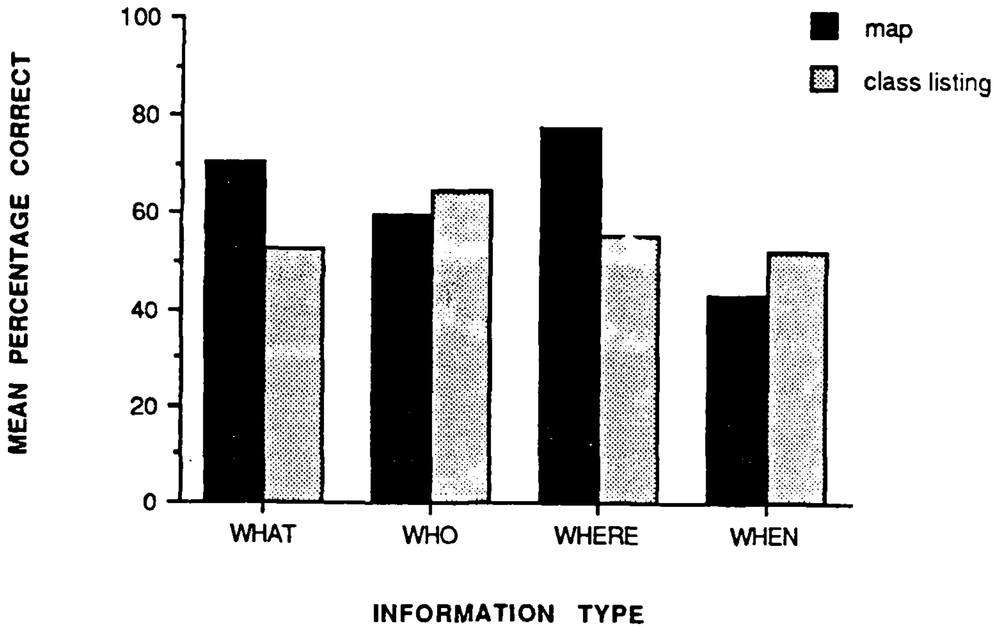


Figure 4. Information Type X Task Type interaction in the questionnaire retest data, Experiment 2.

Separate ANOVAs for each information type revealed that subjects assigned to the map task did significantly better at recalling what information ($M = 70.60\%$) than those assigned to the class listing task ($M = 52.76\%$), [$F(1,28) = 5.11$, $MSe = .33$]. The same appeared true for where information, though the difference between tasks was only marginally significant [$F(1,28) = 3.93$, $MSe = .44$, $p < .06$]. In this case, means for the map and class listing tasks were 77.27% and 55.14%, respectively. Differences for who and when were not significant, though the class listing task tended to show better performance. As in the test data, number of courses was not a significant factor in retest data. Mean percentages correct for 4 and 5 courses were 58.00% ($n = 18$) and 59.50% ($n = 18$), respectively. No interactions with this factor were significant as well.

As for information type, there was a significant main effect [$F(3,84) = 3.26$, $MSe = .22$]. Recall performance from highest to lowest started with where at 64.36%, followed by who, what, and when at 62.28%, 60.19%, and 48.17%, respectively. This was the same ordering found in the test data. However in this case, statistically, where was significantly better than only the when information [$F(1,28) = 6.79$, $MSe = 1.14$]. No interactions with information type were significant (excluding the orienting task interaction cited above). Additional ANOVAs were performed specifically comparing information types from test to retest. All four types of information showed worse performance at retest

than at test, suggesting forgetting had occurred between tests. Mean percentages correct by information type for both the test and retest are provided in Table 10.

Table 10

Mean Percentage Correct by Information Type for Questionnaire Test and Retest

	Information type				
	what	who	where	when	
Test	71.28 *	71.06 *	82.89 *	63.19 *	(72.10)
Retest	60.19	62.28	64.36	48.17	(58.75)

Note. Row means are provided in parentheses.

* Significant differences from test to retest, $p < .05$.

Means for the test reflect only scores from the 36 subjects participating in the retest. Differences among information types from test to retest were significant for what [$F(1,28) = 5.18$, $MSe = .33$], for where [$F(1,28) = 8.41$, $MSe = 1.03$], and for when [$F(1,28) = 4.18$, $MSe = .86$], and marginally significant for who information [$F(1,28) = 3.86$, $MSe = .32$, $p < .06$].

Results of the questionnaire test analysis lend clear support for the superior memory of spatial location information (where) over temporal (when) or item information (what and who) in retention of class schedules. In the retest, the advantage of spatial information was not as evident, possibly because retesting involved the same class information used during initial testing. Analyses of the effectiveness of each information type as a retrieval cue and the type of errors made in recall were performed to explore further the differences among information types.

Analysis of Questionnaire Cue Data

In addition to looking at the four types of information as items to be recalled, they were also examined for their effectiveness as retrieval cues. In the questionnaire test and retest, each subject received three of the twelve possible two-cue combinations, that is, the four information types taken two at a time, the first acting as the primary cue, the other as the secondary cue. Thus, for this analysis, subjects were

grouped in fours such that each set of four subjects (one super-subject) included data from all twelve cue combinations. In total there were twelve super-subjects. Since the three between-subjects factors, training type, task, and number of courses, were not randomly assigned within super-subjects, the analysis performed here was simply a one-way ANOVA with cue type as the single, within-subjects variable. Analyses are reported only for performance after one cue was given (the primary cue). Analyses of performance after two cues did not yield significant differences. Results are first presented for questionnaire test data and then are followed with retest results.

The overall effect of cue type for questionnaire test performance following only the primary cue was significant [$F(3,33) = 4.11$, $MSe = .10$]. Who information was found to be the best cue ($M = 81.97\%$), followed closely by what information ($M = 75.89\%$), with when and where information showing the poorest performance, with means of 65.89% and 62.50%, respectively. Statistically, the who cue was not better than the what cue, but was significantly higher than when and where cues [$F(1,11) = 8.40$, $MSe = .31$, $p < .01$] and [$F(1,11) = 7.93$, $MSe = .46$], respectively. The what cue was not significantly higher than the when cue, but was marginally better than the where cue [$F(1,11) = 4.26$, $MSe = .22$, $p < .06$]. These primary cue data were further examined to see if certain cues were better at eliciting certain types of information, that is, given the who cue, did performance

vary with the type of information recalled? The means for each cue/information type combination are given in Table 11.

Table 11

Mean Percentage Correct Across Information Types as a Function of Cue Type for Questionnaire Test Data

	Information type				
Cue	what	who	where	when	
What	-----	74.83	80.58	72.25	(75.89)
Who	84.75	-----	88.92	72.25	(81.97)
Where	63.92	70.83	-----	52.75	(62.50)
When	58.50	68.25	70.92	-----	(65.89)

Note. Row means are provided in parentheses. There were 12 super-subjects per cell. A cue type was never matched with the same information type.

Of the four cue types, only the where cue showed a significant effect for information type [$F(2,22) = 3.56$, $MSe = .10$]. In this case, the where cue was better at eliciting who information than when information [$F(1,11) = 6.19$, $MSe = .39$, Ms = 70.83% and 52.75%, respectively).

In the analysis of cue type on the questionnaire retest, data were also grouped into 12 observations as described previously. However, because there were only 36 subjects in the retest, not all observations contained data from 4 subjects. Unlike the test data, no significant effect of cue type was found. Despite this, the same pattern across cues appeared. As in the test data, what and who cues appeared to produce the best recall (Ms = 64.58% and 61.14%, respectively). Conversely, the where and when cues appeared less efficient, with recall percentages of 54.47% and 52.14%, respectively. Statistically, the only significant difference was found between the what and where cues, with this difference being marginally reliable [$F(1,11) = 3.69$, $MSe = .12$, $p < .08$].

Table 12 presents the retest data in terms of each cue/information type combination.

Table 12

Mean Percentage Correct Across Information Types by Cue Type for Questionnaire Retest Data

	Information type				
Cue	what	who	where	when	
What	-----	59.00	76.42	58.33	(64.58)
Who	70.83	-----	69.50	43.08	(61.14)
Where	57.75	65.33	-----	40.33	(54.47)
When	51.42	54.92	50.08	-----	(52.14)

Note. Row means are provided in parentheses. There were 12 observations per cell. A cue type was never matched with the same information type.

Looking at differences among information types within each cue type revealed significant differences only for the who cue. For the who cue, there was an overall effect of information type [$F(2,22) = 3.84$, $MSe = .29$]. In this case, both what and where information were recalled marginally better than when information [$F(1,11) = 4.60$, $MSe = .92$, $p < .06$] for what-when, and [$F(1,11) = 3.99$, $MSe = .84$, $p < .07$] for where-when. A similar trend, that is, poorer performance for when information over the other information types, was found given the where cue, but these differences were not significant.

Analysis of Questionnaire Errors

In addition to examining subject accuracy on the class questionnaire test and retest, an analysis of subject errors was also conducted. Performance across information types was looked at first, for differences in mean number of errors made. Because this analysis was largely the complement of the accuracy analysis, it will not be discussed here, but can be found in Appendix G. Presented here is an analysis of errors by error type.

Four classes of errors were identified in the data. These included questions in which (1) no answer was given (2) a wrong answer was given, but the answer was correct for another course (3) a bad answer was given that was not correct for any course on the class schedule, and (4) a partially correct answer was given.

This latter class of errors covered incomplete class titles, misspelled names of professors, and times and locations which were close, but not perfectly accurate. These errors were given partial credit in scoring, but were counted fully here. The mean number of errors (maximum of 3 errors) for the questionnaire test as a function of error type and information type are provided in Table 13.

Table 13

<u>Mean Number of Errors for Questionnaire Test Data as a Function of Error Type and Information Type</u>.

	Information type			
Error Type	what	who	where	when
No answer	0.06	0.19	0.00	0.10
Wrong answer	0.40	0.33	0.27	0.48
Bad answer	0.04	0.00	0.08	0.15
Partial answer	0.40	0.25	0.17	0.06

Separate analyses were conducted for each type of error. For no answer errors, there was a marginally significant main effect of information type [$F(3,120) = 2.60$, $MSe = .30$, $p < .06$]. There were relatively few no answer errors made ($M = 0.35$) with where information having zero no answer errors. Who information had the largest number ($M = 0.19$). This was significantly greater, however, than only where information [$F(1,40) = 8.27$, $MSe = 1.69$]. The wrong answer errors accounted for half of all the errors made ($M = 1.48$), but did not show a significant main effect of information type. For bad answer errors, a main effect of information type was found [$F(3,120) = 2.89$, $MSe = .19$, $M = 0.27$]. Who information, with no bad answer errors, was significantly below where information with a mean of 0.08% [$F(1,40) = 4.00$, $MSe = .33$] and when information, with a mean of 0.15 [$F(1,40) = 6.62$, $MSe = 1.02$]. Finally, in terms of partially correct errors, a main effect of information type was also found [$F(3,120) = 4.68$, $MSe = .95$, $M = 0.88$]. In this case, what information, with a mean of 0.40, was significantly higher than both where information with a mean of 0.17 [$F(1,40) = 4.06$, $MSe = 2.52$], and when information with a mean of 0.06 [$F(1,40) = 15.61$, $MSe = 5.33$].

Results of the error analysis for questionnaire retest data generally followed the same pattern just described for the test data. Table 14 provides the mean error rates for each error type across the four types of information.

Table 14

Mean Number of Errors for Questionnaire Retest Data as a Function of Error Type and Information Type.

	Information type			
Error Type	what	who	where	when
No answer	0.03	0.17	0.08	0.08
Wrong answer	0.47	0.53	0.42	0.86
Bad answer	0.03	0.00	0.19	0.19
Partial answer	0.69	0.33	0.27	0.06

For no answer errors there was no significant main effect of information type. Wrong answer errors, as in the test data, accounted for the greatest number of errors (M = 2.28). There was a significant main effect of information type in this case [F(3,84) = 5.57, MSe = 1.87]. When information yielded the greatest portion of these errors (M = 0.86). This was significantly above the mean number of wrong errors for what [F(1,28) = 14.31, MSe = 7.80, M = 0.47], for who [F(1,28) = 6.00, MSe = 4.41, M = 0.53], and for where information [F(1,28) = 10.80, MSe = 8.93, M = 0.42]. For bad answer errors, the main effect for information type was also significant [F(3,84) = 3.52, MSe = .30, M = 0.42]. In this case, significant differences were found between who information, with zero bad answer errors, and where information with a mean of 0.19 [F(1,28) = 8.58, MSe = 1.19], and when information, also with a mean of 0.19 [F(1,28) = 4.47, MSe = .85]. This was the same result found in the test data. Finally, for partially correct errors, a main effect for information type was also found [F(3,84) = 8.99, MSe = 2.23, M = 1.25]. What information, with a mean of 0.65, clearly had more of these errors than who information with a mean of 0.33 [F(1,28) = 9.23, MSe = 4.06], where with a mean of 0.17 [F(1,28) = 9.03, MSe = 8.21], and when information with a mean of 0.06, just as found in test data.

Several significant observations come from the analyses of cue efficiency and recall error data. First, despite the

possible advantage of spatial location (where) information at recall, spatial location information along with temporal (when) information served as relatively poor retrieval cues. In contrast, item information (who and what) were much better as recall cues. In terms of recall errors, results suggest that for what information, more than for the other types of information, subjects could remember part, but not all of the item to be recalled. This was especially true on the questionnaire retest. Further, performance on when information during the retest reflected a disproportionately large number of confusion errors (wrong answers) with other times in the class schedule. This finding suggests that, over time, temporal information (class start times) is subject to large interference effects.

Analysis of the Map and Class Listing Test Data

During testing, each subject completed the questionnaire discussed above and then completed both the same map and class listing recall tasks used in the training procedure. The map and class listing test, as they are referred to here, were administered in a counterbalanced manner across subjects. Scoring of the two tests was identical to that used in the questionnaire data. Subjects were assigned a percentage correct for each of the four types of information in their schedule, the what, who, where, and when of each course. The primary interest in these data was differences across information types and their

interactions with the three between-subjects factors, training type, orienting task, and the number of courses. Before presenting these results, it is important to note several significant main effects for these other variables.

Table 15 provides the mean percentages correct across information types by training, task, and number of courses for the map test data.

Table 15

Mean Percentage Correct for Map Test Data Across Information Types as a Function of Training, Task, and Number of Courses

	Information type *				
	what	who	where	when	
Training *					
Massed (n = 22)	85.00	68.59	89.91	83.32	(81.70)
Spaced (n = 24)	97.92	94.75	99.50	97.29	(97.36)
Task					
Map (n = 24)	90.58	76.33	95.00	93.33	(88.81)
Class Listing (n = 22)	93.00	88.68	94.82	87.64	(91.03)
Courses					
4 (n = 23)	90.87	82.74	94.61	92.96	(90.29)
5 (n = 23)	92.61	81.74	95.22	88.26	(89.46)
	(91.74)	(82.24)	(94.91)	(90.61)	(89.87)

Note. Column and row means are provided in parentheses.

* Significant main effect at $p < .05$.

The mean overall performance on the map test was 89.87% ($n = 46$, 2 subjects neglected to fill in the campus map and were not used in the analysis). Of the three between-subjects factors, only training type had a significant main effect [$F(1,38) = 25.36$, $MSe = 1.12$]. Mean percentage correct for spaced subjects was 97.36% and 81.70% for massed training subjects. Surprisingly, there was no main effect of task [$F(1,38) = .23$, $MSe = .01$, $p < .63$]. Subjects who trained with the map task did not do better on the map test than those who trained using the schedule task (Ms = 88.81% and 91.03% for the map and class listing task, respectively). Number of courses was also not a significant factor in the map test data.

There was a main effect of information type in the map test [$F(3,114) = 7.35$, $MSe = .13$]. Mean percentages for where, what, when, and who, in order, were 94.91%, 91.74%, 90.61%, and 82.24%. Single degree of freedom tests revealed that performance for the where information was significantly better than that for when [$F(1,38) = 4.71$, $MSe = .10$] and who [$F(1,38) = 24.49$, $MSe = .73$]. Though performance for where was numerically higher than that for what, this difference was not statistically significant, possibly because performance was so close to the ceiling. The map test data also revealed an Information X Training Type interaction [$F(3,114) = 2.99$, $MSe = .05$]. Subjects in the massed training condition performed significantly worse on who information than those receiving spaced training. Here again, it is likely that

differences among information types were not apparent for spaced trained subjects because performance was at the ceiling (see Figure 5).

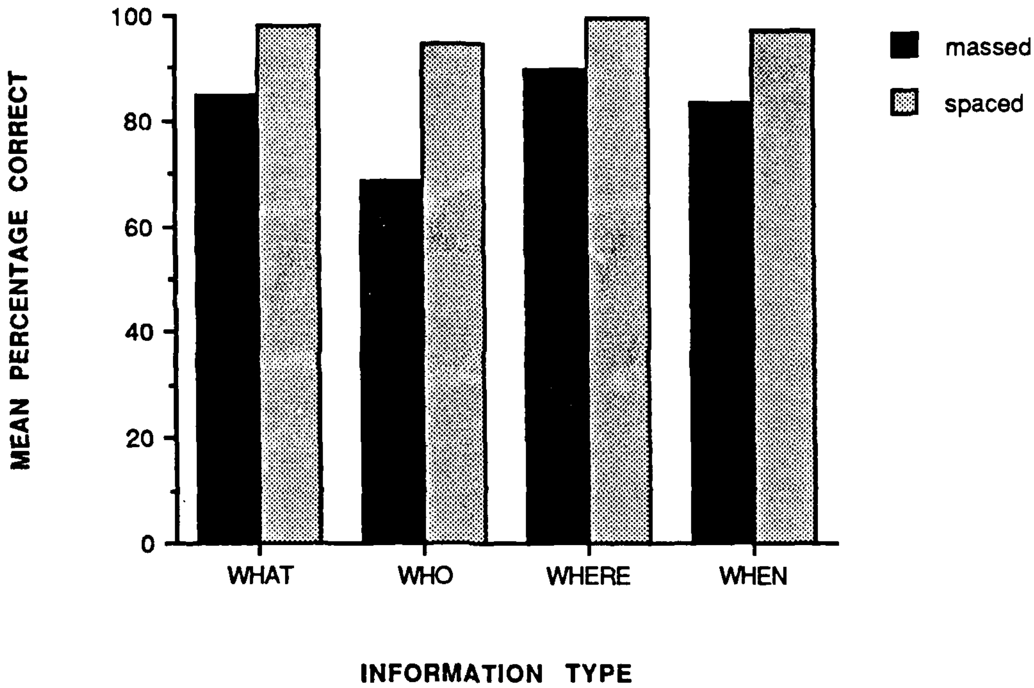


Figure 5. Information Type X Training Type interaction in the map test data, Experiment 2.

Finally, it was found that information type differences on the map test also varied with the assigned task during training [$F(3,114) = 3.58$, $MSe = .06$]. As in the Information X Training Type interaction, this interaction appeared due to differences in performance for who information. Subjects assigned to the map orienting task performed more poorly on who information than those using the class listing task. This difference was not evident in the training data where subjects in both task groups performed equally at the conclusion of training. The difference in this case was marginally significant [$(1,38) = 3.33$, $MSe = .14$, $p < .08$], (see Figure 6).

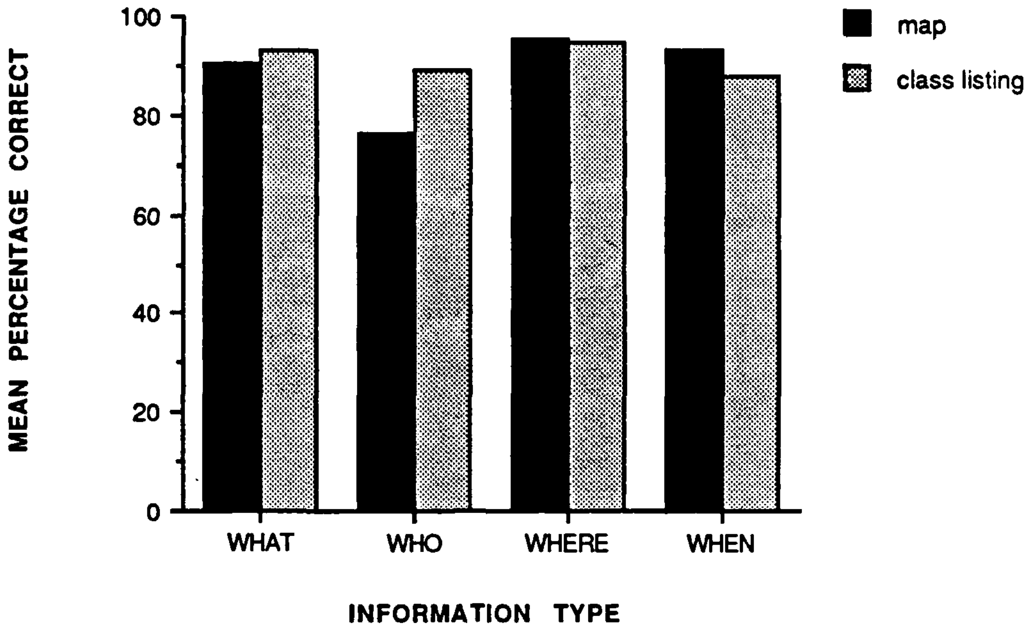


Figure 6. Information Type X Task Type interaction in the map test data, Experiment 2.

Table 16 provides the mean percentages correct across information types by training, task, and number of courses for the class listing test data.

Table 16

Mean Percentage Correct for Class Listing Test Data Across Information Types as a Function of Training, Task, and Number of Courses

	Information type *				
	what	who	where	when	
Training *					
Massed	83.13	73.08	71.21	65.13	(73.14)
Spaced	97.42	94.75	89.21	95.42	(94.20)
Task *					
Map	87.04	77.17	73.08	77.50	(78.70)
Class Listing	93.50	90.67	87.33	83.04	(88.64)
Courses					
4	89.71	84.50	80.83	83.88	(84.73)
5	90.83	83.33	79.58	76.67	(82.60)
	(90.27)	(83.92)	(80.21)	(80.27)	(83.67)

Note. Number of subjects per cell was 24. Column and row means are provided in parentheses.

* Significant main effect at $p < .05$.

Mean overall performance on the class listing test was 83.67%, $n = 48$. In terms of main effects, both training and task type effects were significant [$F(1,40) = 23.08$, $MSe = 2.13$] for training type and [$F(1,40) = 5.14$, $MSe = .47$] for task type. As in the map test data, spaced training subjects showed better performance over those receiving massed training ($Ms = 94.20\%$ and 73.14%, respectively). Unlike the map test data, however, subjects who trained using the class listing task did better on the class listing test than those who trained using the map task ($Ms = 88.64\%$ and 78.70%, respectively). Number of courses was not a significant factor in the class listing test data.

A significant main effect of information type was also found in the class listing test [$F(1,40) = 4.43$, $MSe = .11$]. However, the pattern of differences was not the same as found in the map test data. Performance for where information ($M = 80.21\%$) was no different than that for who ($M = 83.92\%$) and when ($M = 80.27\%$) information. For the class listing test data, what information ($M = 90.27\%$) showed the best recall. Performance for what information was marginally better than the who information [$F(1,40) = 3.75$, $MSe = .19$, $p < .06$], and significantly better than where [$F(1,40) = 18.55$, $MSe = .49$], and when information [$F(1,40) = 9.62$, $MSe = .48$]. In addition, unlike in the map test data, information type was not found to interact with training or task type, nor was there an interaction with number of courses.

Analysis of Map and Class Listing Retest Data

Following completion of the questionnaire retest, subjects were once again administered the map and class listing tests. The mean overall performance on the map retest was 81.15%, (n = 34, two subjects neglected to fill in the campus map). This was significantly below the overall performance on the initial map test of 88.63% for the same 34 subjects [$F(1,26) = 15.19$, $MSe = .15$]. Table 17 provides mean percentages correct across information types by training, task, and number of courses for map retest data.

Table 17

Mean Percentage Correct for Map Retest Data Across Information Types as a Function of Training, Task, and Number of Courses

	Information type *				
	what	who	where	when	
Training *					
Massed (n = 17)	81.24	60.12	88.06	66.76	(74.04)
Spaced (n = 17)	88.47	81.29	96.82	86.47	(88.26)
Task					
Map (n = 15)	88.27	65.13	94.53	81.33	(82.32)
Class Listing (n = 19)	82.16	75.11	90.79	72.89	(80.24)
Courses					
4 (n = 17)	83.24	76.71	94.29	77.94	(83.04)
5 (n = 17)	86.47	64.71	90.59	75.29	(79.26)
	(84.85)	(70.71)	(92.44)	(76.62)	(81.15)

Note. Column and row means are provided in parentheses.

* Significant main effect at $p < .05$.

As with the map test data, performance on the map retest revealed a main effect of training type [$F(1,26) = 7.14$, $MSe = .38$]. Spaced training ($M = 88.26\%$, $n = 17$) was shown to yield better performance than massed training ($M = 74.04\%$, $n = 17$). Unlike in the map test, however, training type did not interact significantly with information type. Massed training subjects did worse on all four types of information, though differences did appear greatest for who and when information, as in the test data. Also similar to test results, task type and number of courses did not show significant main effects. Furthermore, there were no significant interactions with these factors, even though there was a significant Task X Information Type interaction in the test data.

The most important results of the map retest data come from the analysis of information type. There was a main effect of information type [$F(3,78) = 7.33$, $MSe = .28$], but more critically, where information was better recalled than any of the other three types of information. Mean percentages for where, what, when, and who, in order, were 92.44%, 84.85%, 76.62%, and 70.71%. All three other means were significantly below the mean for where, [$F(1,26) = 9.20$, $MSe = .19$] for where versus what, [$F(1,26) = 9.87$, $MSe = .70$] for where versus when, and [$F(1,26) = 18.89$, $MSe = 1.49$] for where versus who. Furthermore, comparisons made by information type from test to retest showed significant forgetting between tests for what information

[F(1,26) = 11.44, MSe = .14], for when [F(1,26) = 5.13, MSe = .33], and for who information [F(1,26) = 8.22, MSe = .22, $p < .008$]. In contrast, where information showed no significant loss between test and retest, despite the five week delay. Mean percentages correct for the test-retest comparisons can be found in Table 18.

Table 18

Mean Percentage Correct by Information Type for Map Test and Retest

	Information type				
	what	who	where	when	
Test	92.17 *	79.32 *	95.18	87.88 *	(88.63)
Retest	84.85	70.71	92.44	76.62	(81.15)

Note. Row means are provided in parentheses.

* Significant differences from test to retest, $p < .05$.

These results suggest a certain durability in memory for the spatial location of courses. The same cannot be said for memory for the names of the buildings where courses were held, as is shown next in the class listing retest data.

The mean overall performance on the class listing retest was 70.61% ($n = 35$, one subject neglected to fill in the when information on the test and was excluded from the analysis). This was significantly below the overall performance on the initial class listing test of 82.55% for the same 35 subjects [$F(1,27) = 20.31$, $MSe = .42$]. Table 19 provides mean percentages correct across information types by training, task, and number of courses for class listing retest data.

Table 19

Mean Percentage Correct for Class Listing Retest Data Across Information Types as a Function of Training, Task, and Number of Courses

	Information type *				
	what	who	where	when	
Training *					
Massed (n = 19)	80.47	63.95	52.89	44.21	(60.38)
Spaced (n = 16)	87.13	78.25	82.69	83.00	(82.77)
Task					
Map (n = 14)	86.71	59.79	66.50	60.57	(68.39)
Class Listing (n = 21)	81.38	77.62	66.52	62.86	(72.10)
Courses					
4 (n = 18)	83.50	75.94	65.44	60.44	(71.33)
5 (n = 17)	83.53	64.71	67.65	63.53	(69.85)
	(83.51)	(70.49)	(66.51)	(61.94)	(70.61)

Note. Column and row means are provided in parentheses.

* Significant main effect at $p < .05$.

Consistent with the test data, the class listing retest results revealed a significant main effect of training type [$F(1,27) = 16.26$, $MSe = 1.52$, $p < .0004$]. As in all the previous analyses, spaced training ($M = 82.77\%$, $n = 16$) produced significantly better performance than massed training ($M = 60.38\%$, $n = 19$). There was also a significant Information X Training Type interaction [$F(3,81) = 2.90$, $MSe = .16$]. As can be seen in Figure 7, the advantage of spaced over massed training was less pronounced for what and who information than for where and when.

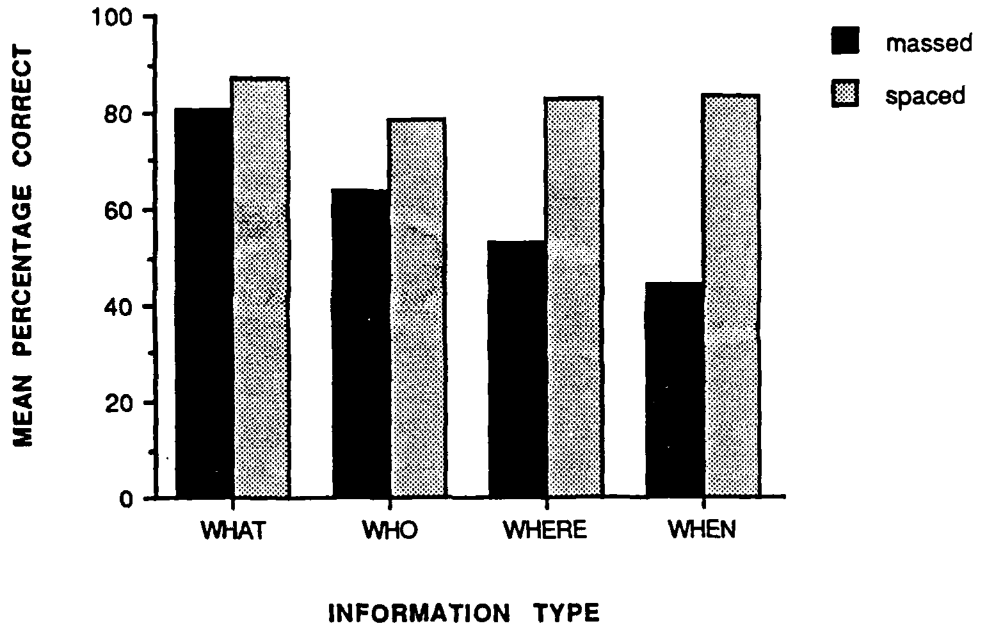


Figure 7. Information Type X Training Type interaction in the class listing retest data, Experiment 2.

In the case of what and who, the training type differences were not significant. For where and when, there was a significant effect for training type [$F(1,27) = 8.74$, $MSe = .68$, $Ms = 82.69\%$ for spaced and 52.89% for massed training for where information], and [$F(1,27) = 12.50$, $MSe = 1.11$, $Ms = 83.00\%$ for spaced and 44.21% for massed training for when information]. In addition, unlike the test results, there was no significant main effect found for task type, nor was there a significant effect for number of courses. Neither of these two factors yielded any significant interactions.

As for information type, class listing test and retest results were very consistent. Like the test results, the main effect of information type was significant [$F(3,81) = 4.41$, $MSe = .24$]. Also consistent was the superior performance of what information ($M = 83.51\%$) followed by who, where, and when ($Ms = 70.49\%$, 66.51%, and 61.94%, respectively). Performance for where information was not significantly different from who or when and was significantly below performance for what information [$F(1,27) = 12.34$, $MSe = .81$]. This suggests that recalling a building's name is not the same as recalling its spatial location on a map. Furthermore, in the class listing data there was a significant decrease in performance from test to retest for where information [$F(1,27) = 7.74$, $MSe = .64$], unlike the lack of forgetting for where information in the map test-retest data. In addition to the decline in performance for where information, what, who, and

when information also showed a significant decrease, for what [$F(1,27) = 7.15$, $MSe = .15$], for who [$F(1,27) = 10.74$, $MSe = .36$], and for when [$F(1,27) = 6.60$, $MSe = .66$]. Mean percentages correct for the test-retest comparisons can be found in Table 20.

Table 20

Mean Percentage Correct by Information Type for Class Listing Test and Retest

	Information type				
	what	who	where	when	
Test	90.34 *	80.91 *	81.14 *	77.80 *	(82.55)
Retest	83.51	70.49	66.51	61.94	(70.61)

Note. Row means are provided in parentheses.

* Significant differences from test to retest, $p < .05$.

Results of the map and class listing tests provide further support and clarification of the role of spatial information in class schedule recall. Two observations are particularly important. First was the striking finding that in the map test and retest, not only did where information (in this case, the spatial location of classes on a map) show superior performance over the other types of information, but there was also no loss of information over the five week retention interval. What, who, and when information all showed significant losses. The second critical finding was that in the class listing test, where information (in this case, the name of the building where classes were held) showed no superior performance over the others and showed significant loss at retest, five weeks later. This suggests the advantage of where information can be attributed to its truly spatial aspects.

Analysis of Training Data

The main concern in this analysis was that at the conclusion of training, subjects had learned the four types of information in their schedules equally well. This was found to be the case. Analysis of performance during the ninth and last training trial revealed no significant main effect of information type. Mean percentages correct for what, who, where, and when information were 99.58%, 97.96%, 98.43%, and 100.00%, respectively. But this is not to say that performance for each type of information

progressed equally across training trials. The mean percentages correct for each information type across the nine training trials are plotted in Figure 8.

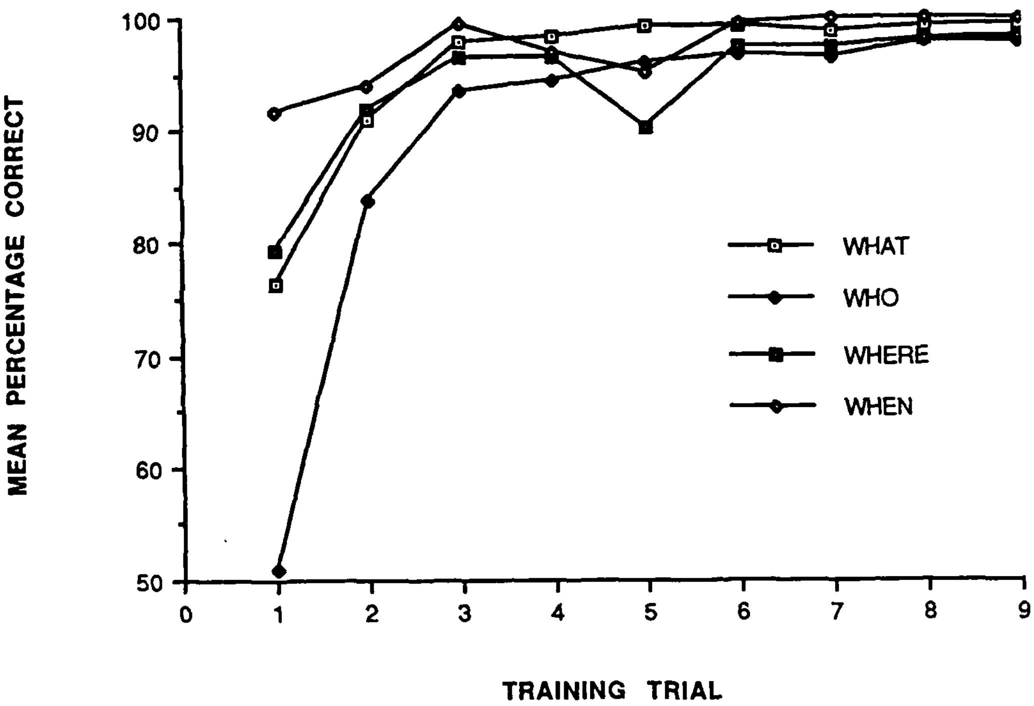


Figure 8. Mean percentage correct across training trials by Information Type.

There was a significant main effect of information type in each of the first three trials, during trial 5 (when subjects switched tasks), and again during trial 7. No significant main effect of information type was found in trials 4, 6, 8, and 9. Each of the significant main effects will be discussed momentarily, but before doing so, results collapsing across training trials will be presented.

Performance during training was analyzed using the same 2 X 2 X 2 X 4 mixed design followed thus far. In this case, the 4-level within-subjects factor included mean performance for each type of information across the 9 training trials. The mean overall performance for training data was 94.33%, (n = 44, 2 subjects arrived late and missed trial 1, 2 other subjects neglected to fill in the campus map, one on trial 6, the other on trial 9). Mean percentages across information types by training, task, and number of courses are provided in Table 21.

Table 21

Mean Percentage Correct for Training Data Across Information Types as a Function of Training, Task, and Number of Courses

	Information type *				
	what	who	where	when	
Training					
Massed (n = 20)	95.29	90.14	93.49	96.68	(93.90)
Spaced (n = 24)	96.32	89.50	94.49	98.42	(94.68)
Task					
Map (n = 22)	95.27	87.25	92.62	97.86	(93.25)
Class Listing (n = 22)	96.43	92.34	95.45	97.40	(95.41)
Courses					
4 (n = 23)	96.66	91.10	94.87	98.31	(95.24)
5 (n = 21)	94.97	88.36	93.12	96.88	(93.33)
	(95.85)	(89.79)	(94.04)	(97.63)	(94.33)

Note. Column and row means are provided in parentheses.

* Significant main effect at $p < .05$.

The three between-subjects factors and their interactions showed no significant main effects. This is not surprising since performance was already near the ceiling by the end of trial three. In contrast, the main effect of information type was significant [$F(3,108) = 13.19$, $MSe = .05$]. Mean percentages for what, who, where and when were, in order, 95.85%, 89.97%, 94.04%, and 97.63%. Differences here likely reflect two influences, the relatively slow learning of who information and rapid learning of when information, especially during the first three training trials. For example, on average, it took subjects 3.94 training trials to achieve the first perfect score (fully correct on all 4 or 5 courses) for who information. In contrast, it took an average of only 1.38 trials for the first perfect score on when information. The averages for what and where information were 2.40 and 2.29 trials, respectively. These differences will be presented more fully in the forthcoming trial-by-trial analysis.

The mean overall performance for trial 1 was 74.58% ($n = 46$, 2 subjects arrived late and missed this trial). Only number of courses proved to be a significant between-subjects factor [$F(1,38) = 5.26$, $MSe = .57$]. As might be expected, subjects with 4 courses ($M = 80.02\%$, $n = 23$) did better on trial 1 than those with 5 courses in their schedules ($M = 69.13\%$, $n = 23$). In terms of information type, the largest performance differences were found for training trial 1 [$F(3,114) = 32.57$, $MSe = 1.33$]. During this trial, subjects recalled who information ($M = 51.04\%$)

worse than either what, where, or when information (Ms, in order, = 76.35%, 79.32%, and 91.59%) with F ratios of [F(1,38) = 55.54, MSe = 2.95], [F(1,38) = 25.86, MSe = 3.70], and [F(1,38) = 96.25, MSe = 7.51], respectively. Conversely, when information was recalled better than any of the other three types of information, for what [F(1,38) = 22.45, MSe = 1.05], for who [F(1,38) = 96.25, MSe =7.51], and for where [F(1,38) = 8.36, MSe = .67]. It is important to keep in mind that subjects were instructed to study their fictitious schedule just as if it were their actual schedule. Thus, results of trial 1 suggest that subjects first learned (and perhaps more easily) when their classes occurred followed by where classes were held, what they were titled, and finally, who was the instructor.

The mean overall performance for trial 2 was 90.21%, n = 48. No significant between-subjects effects were found. For information type, results from trial 2 showed smaller performance differences than in trial 1, though there continued to be a significant main effect [F(3,120) = 4.31, MSe = .09]. For this trial, the what, where and when information types were recalled equally well (Ms, in order, = 91.06%, 91.98%, 93.96%). However, performance for who information (M = 83.83%) was still significantly below each of the other three types, for what [F(1,40) = 4.56, MSe = .258], for where [F(1,40) = 8.44, MSe = .32], and for when [F(1,40) = 12.20, MSe = .49]. By the third training trial, overall performance was 96.90%, n = 48. No

significant between-subjects effects were found. However, a significant main effect of information type was still found [$F(3,120) = 5.93$, $MSe = .030$]. Means for what, who, where, and when information were 97.83%, 93.58%, 96.58%, and 99.58%, respectively. Performance for who information was still lowest, but significantly below only what and when information [$F(1,40) = 10.08$, $MSe = .09$] and [$F(1,40) = 18.40$, $MSe = .17$], respectively. In addition, subjects were performing almost perfectly on the when information, leaving performance for what, who and where information types significantly below [$F(1,40) = 5.76$, $MSe = .014$], [$F(1,40) = 18.40$, $MSe = .17$], and [$F(1,40) = 4.14$, $MSe = .04$], respectively.

The main effect of information type in trial 4 was not significant nor were there significant between-subjects effects. Mean overall performance was 96.55%, $n = 48$. This was not the case in trial 5 where overall performance appeared to drop slightly ($M = 95.22\%$, $n = 48$), though not significantly. During this trial subjects were asked to practice recall using the alternate orienting task. There was some evidence that this decrease was related to training type. There was a marginally significant Information X Training Type interaction [$F(3,120) = 2.54$, $MSe = .03$], as shown in Figure 9.

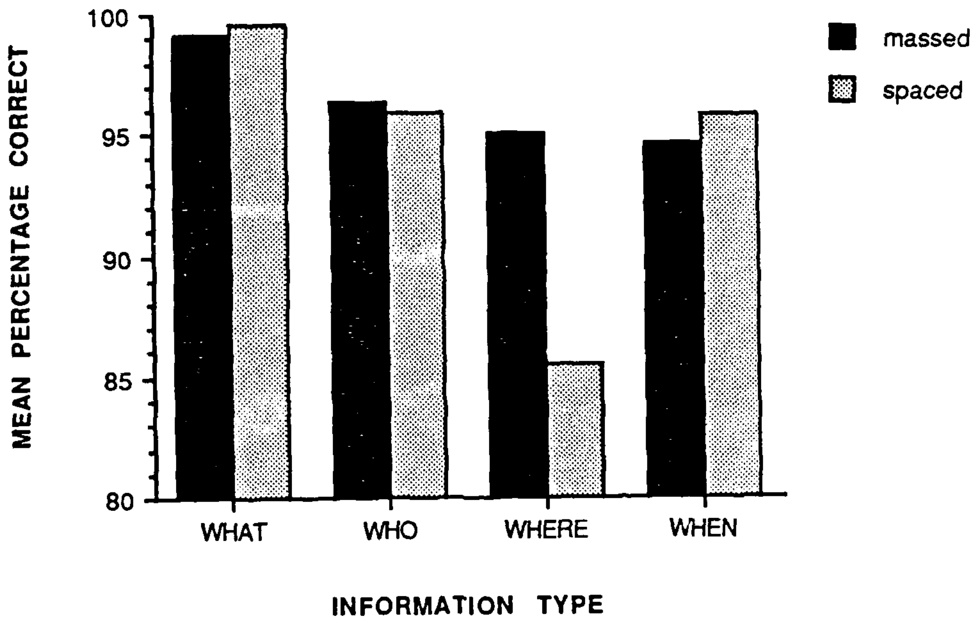


Figure 9. Information Type X Training Type interaction in the training data on trial 5, Experiment 2.

Subjects receiving massed training performed significantly better on where information (M = 95.08%) than those given spaced training (M = 85.45%), [$F(1,40) = 4.49$, $MSe = .11$].

Finally, from trials 6 through 9 there appeared to be little difference in performance on the four information types. A small main effect for information type was found for trial 7 [$F(3,120) = 2.65$, $MSe = .01$] and appeared to be due to performance on the when information reaching the ceiling at a mean percentage correct of 100%. In addition, in trial 7 who information continued to lag slightly behind the other information types, being significantly below the what and when information types [$F(1,40) = 3.98$, $MSe = .02$] and [$F(1,40) = 11.24$, $MSe = .06$], respectively.

In sum, the training data revealed the following observations. Subject performance improved rapidly during the first three training trials. Differences in performance during these trials suggested that subjects learned the when information first and perhaps more easily and gave last preference and perhaps had more difficulty with who information. Overall, performance continued to improve in trials 4 through 9, though at a much slower rate. There were a few exceptions to this trend. Performance for where information decreased in trial 5 with the switch in orienting tasks, but only for spaced trained subjects. Another notable trend was overall performance on when information. It was the only type of information in which

performance reached 100% correct, with performance remaining perfect for trials 7, 8 and 9. Finally, given our primary interest in the where information, the training data did not suggest a training explanation for the where information advantage in the delayed recall data.

Discussion

Results of this experiment, as in Experiment 1, lend support for the superior retention in long-term memory of where information, relative to what, who, and when information. More specifically, results suggest the where advantage holds only for identification of map locations and not for recall of names of locations. A number of factors may have contributed to this advantage or lack of advantage for where information. In Experiment 1, two factors were considered; the benefits provided directly by the campus map and the possible contributions of spatial knowledge. Within these two factors were considered five specific accounts of the where recall advantage. Each of these accounts or hypotheses will be considered in light of new evidence from this experiment. To best evaluate the merit of each proposed account, it is necessary to consider them in light of the specific procedures and materials used in this experiment.

To review briefly, testing in the questionnaire followed the same cued recall format used in Experiment 1. Subjects were cued with one of the four types of information and then were

questioned about the other three. Questions required writing out the course title (what), the professor's name (who), the class start time (when), and identifying on a map the location of a course (where). For what, who, and when information, responses were provided one at a time, in an established order. This format was very different from the format used in training. Training recall sessions used a less structured format which allowed subjects more freedom in the order information was recalled. For example, subjects could list all the course titles first and then all the names of professors, etc., or they could complete all four types of information, course by course. For where information, however, the only difference between training and testing was that in training all courses were located on one map. In testing, locations were marked one course at a time on separate maps. The same map was used in training and testing, though in testing all verbal labels were removed.

Testing in the map and class listing tests matched training procedures much more closely. Both tests used the same open format found in training which required subjects to recall all their course information without specifying recall order. The map test involved listing the names of courses and professors, the order of their courses during the school week (when information), and then locating on a single campus map all the course locations, without reference to specific courses or order. The class listing test involved listing the names of courses and

professors as well as the names of buildings and course times. No map was used in this test.

Each of the five hypotheses proposed in Experiment 1 will be examined in terms of the major results of this experiment. These results include a) the finding of a where advantage for the questionnaire test, but no significant advantage at retest, b) the finding of a where advantage for both the map test and retest, and c) the lack of a where advantage in both the class listing test and retest. Following this evaluation, the hypotheses will be considered in light of more specific results involving training and task effects.

The first two hypotheses considered are the multiple cues and recognition hypotheses. Both suggest that the where advantage was due to a direct contribution of the map during testing. The multiple cues hypothesis simply holds that, unlike in the what, who, and when tasks, where information was cued by multiple cues because the map was presented during recall. This hypothesis accounts for the where advantage in the questionnaire test, but does not easily explain the significant reduction in the where advantage at retest. By this account, one would expect an across-the-board decrease in performance for all four types of information, but a significant where advantage is still expected since the map continued to be available. Further evidence against this hypothesis is the fact that in the map test, the what, who, and when tasks received multiple cuing as well.

Subjects were asked to recall information using the same forms used in training. These recall forms presented the department of the course, the course number and the column labels, all studied in training. In spite of this additional cuing, where information continued to be recalled better. In addition, evidence from previous research showing a where advantage when cuing conditions were equal is also contrary to this hypothesis (i.e., Pezdek, et al., 1986).

The recognition hypothesis holds that, because the map provided the subjects with all the possible alternative locations, the where task was really a recognition test, not a recall test. Consequently, performance on the where task was expected to be better than the cued recall performance required in the what, who, and when tasks. This hypothesis has the same difficulties as the multiple cues hypothesis. It cannot account for why a significant advantage for where was not found in the questionnaire retest yet a significant where advantage was found in the map retest. Class listing test and retest results do provide support for this hypothesis. The where task in this case was clearly a cued recall test since only building names were required and no map was used.

The next three hypotheses, those which suggest that some aspect of spatial knowledge was responsible for the where advantage, also showed mixed success in accounting for the present results. The spatial procedures hypothesis can be

evaluated from two viewpoints. Assuming a reinstatement of procedures viewpoint (Kolers & Roediger, 1984), all performance can be described within the framework of acquired skills or procedures. It is assumed that learning involves the acquisition of procedures that are specific to the activities involved in the learning task. Thus, the success of memory performance is directly tied to the degree the activities of the memory task match those of the acquisition task. The guiding principle is that procedures acquired during learning can be reinstated at test. The greater the match between training and test, the greater potential for reinstatement and the better the performance. The advantage for where information in questionnaire test performance is easily explained by this account. The where task yielded the best recall because it had the greatest degree of training-test similarity. In locating classes, subjects used the same map as used in training and were thus able to reinstate previously learned procedures at test. As previously discussed, the where task was in some respects different from training. Subjects located one course on each map as opposed to all courses on the same map. On the other hand, the what, who, and when tasks were more different from their respective training tasks. In training, information was entered on a blank form with only column and row headings. No specific order of recall was required. In the testing task, subjects read a question and filled in a blank space, one response at a time,

in a prescribed order. Because of the greater training-test similarity in the where task, subjects were able to reinstate more previously learned procedures and thus produce better recall.

At retest, questionnaire performance did not show a significant advantage for where information. Here an assumption is necessary to account for this lack of advantage. It is assumed that familiar procedures are not as easily reinstated as time between training and testing (the retention interval) lengthens. In the case of the questionnaire retest, the reinstatement advantage for where was lost after six weeks. The reinstatement of procedures account has greatest difficulty explaining the where advantage in the map test and retest. In this case, all four tasks were similar to the training tasks. Yet the where advantage was still found. The reinstatement account cannot explain this advantage at test and its persistence at retest. Results of the class listing test and retest are more easily explained. As with the map test, training-test similarity was high for all four tasks. The lack of a where advantage was expected. In sum, the reinstatement of procedures framework can account for many, but not all of the major results of this experiment.

The same conclusion holds for the proceduralization of knowledge account described by Anderson (1982). In this view, knowledge is seen as having both a declarative and procedural

form. Through extended practice, a behavior such as going to and from classes can become proceduralized. In the process, the declarative components of the behavior, the names of buildings, courses, etc., are forgotten. This account had a certain intuitive appeal in Experiment 1 where schedules were learned naturally over the course of a semester. In Experiment 2, however, subjects did not actually proceed to and from classes. Still, where information was found to have an advantage on the questionnaire and map test. It could be assumed that subjects covertly proceduralized class location information and not the more declarative what, who, and when information. But this does not explain why a where advantage was found for the map retest and not for the questionnaire retest. In addition, subjects were trained to the same level of performance for all four types of information. It is therefore difficult to argue that the declarative information was ignored or forgotten in the process of proceduralizing.

The second spatial hypothesis considered the role of visual images. This hypothesis holds that the map facilitated visual imaging of building locations and other features. Thus, in completing the where task, subjects made use of both their visual knowledge of the campus and their knowledge of building names. By having both visual and verbal information available, they did better on the where task. The what, who, and when tasks are assumed to rely predominately on verbally encoded information at

recall. This explanation easily accounts for the advantage for where information in the questionnaire test. Explaining the lack of where advantage at retest was not as simple. It could be assumed that at retest, only a visual code was available for recalling where information. That is, subjects had forgotten building names, but retained their visual knowledge of building locations. Since all four tasks relied predominately on single codes, albeit different codes, performance was expected to be the same. But this account begs the question of why the largely verbal what, who, and when tasks did not show a commensurate drop in performance at retest, leaving where task recall still at the top. Finally, as with the proceduralization hypothesis, the best test of this account was to explain why there was no where advantage for the questionnaire retest while on the other hand, after five weeks, there was a where advantage for the map retest. A purely visual encoding explanation cannot account for this disparity.

The third spatial memory account, the spatial retention hypothesis, also has difficulty accounting for the questionnaire retest/map retest disparity. By this account, the where task showed a recall advantage because it relied on the use of spatial relations information. This account depends on the assumption that spatial relations information has unique processing characteristics, that is, it is better retained than the nonspatial knowledge found in the other tasks. The where

advantage in the questionnaire test results is taken as an example of better retention of spatial relations information. The lack of a where advantage for the questionnaire retest is attributed to the longer retention delay. Relational information was assumed useful within one week of training, but after five more weeks, the information was assumed no longer available. Again, the difficulty is in explaining why five weeks after the map test, subjects continued to show a where advantage in the map retest.

In summary, the multiple cues and recognition test hypotheses had difficulty explaining changing performance from questionnaire test to retest, and could not easily account for the lack of where advantage in the questionnaire retest while still finding an advantage in the map retest. The three spatial knowledge hypothesis found greater support, but no single hypothesis could explain all results without additional assumptions.

Before drawing any conclusions about the value of these explanations in accounting for the where advantage, it is important to review findings for the between-subjects variables in this experiment, specifically, training and task effects. Number of courses was not considered since it did not show a significant effect in any of the tests.

Spaced training yielded better performance than did massed training in almost all cases. The one exception was in the

questionnaire test. In this case, spaced training led to superior performance for what, who, and when information, but not for where. That is, both massed and spaced trained subjects performed equally on where information recall. As suggested earlier in the chapter, this interaction is likely due to performance for where information being on the ceiling. Thus, differences in training type provide little further help in identifying the locus of the where advantage. Some other factor, i.e., encoding variability (Glenberg, 1979) is likely responsible for the spacing effect. Results of the training task analysis are more revealing, favoring the procedural reinstatement hypothesis.

Task assignment directly determined two things: a) how much practice subjects received in recalling course information, and b) in what format the practice took. Though all subjects had the same amount of study time, there were differences in what was practiced and how much. Reviewing briefly, subjects participated in nine study-then-recall trials during the training phase of the experiment. Subjects assigned to the map task during training performed recall using the map test format on eight trials and on one trial used the class listing test format. The reverse was true for those subjects assigned to the class listing task. On eight trials they practiced recall using the class listing test format and on one trial, the map test format. It would seem reasonable to expect that the task most practiced would yield the

best performance at test. To some extent, this expectation held up in the data. Those assigned to the class listing task did better overall on the class listing test. This result supports the procedural reinstatement hypothesis. The procedures learned in training could be reinstated at test. In contrast, there was no overall task effect for the map test. Even with the extra recall practice received by those in the map task, this added practice did not lead to better recall performance. On the surface, this result does not appear consistent with the procedural reinstatement hypothesis. However, it might be argued that subjects in both study tasks proceduralized spatial locations directly from the study map, that is, mentally, without any overt motor practice. Consequently, the effect of practicing the reinstatement of this procedural information was reduced. There was evidence for a task difference at retest. This may be where the benefit of extra practice was realized. Though differences did not quite reach statistical significance, map trained subjects did show better performance on what, where and when information than that for those in the class listing task on the map retest.

Finally, in the questionnaire test, there was a marginal overall effect for task, favoring the class listing task subjects for all types of information. This result demonstrates that practice on the class listing task helped questionnaire performance more than did practice on the map test. In terms of

procedural reinstatement, it might be said the class listing task had greater procedural correspondence with the questionnaire test than did the map task. The result was the class listing task advantage in test.

In summary, results of task type analysis lends greatest support to the procedural reinstatement account of the where advantage. Other major results also suggest that procedural reinstatement account was implicated. On the other hand, the spatial advantage in the map test when all four types of information involved reinstatement of procedures provided critical contrary evidence. In the next chapter, results of both experiments will be considered in terms of a combined explanation of the where advantage, involving proceduralization and factors unique to spatial information.

CHAPTER IV

GENERAL DISCUSSION

The primary goal of this research was to examine the long-term retention characteristics of spatial knowledge. A natural learning paradigm was used which allowed the study of retention across years without having to wait years for forgetting to occur. Memories for class schedule information were studied in terms of their temporal, spatial, and item components. A natural memory research design was used that allowed objective determination of the information originally studied. Furthermore, using a cued recall questionnaire prepared by the experimenter avoided recording bias and restriction to a single subject design. Thus, Experiment 1 examined naturally learned memories in a manner which overcame many of the common drawbacks of natural memory research. Results are important for this reason alone. But this research had as its secondary goal the comparison of natural memory results with results found for the same types of information learned in the laboratory (Experiment 2). Where information recall (the spatial component) was found to be superior in both experiments.

In Experiment 1, where information showed a strong advantage

across even the longest retention intervals, and did not show significant forgetting until after the second year. Recall for what, who, and when information (the item and temporal components) was below that for where information. Experiment 2 also demonstrated superior retention for where information. In this case the where advantage varied with the experimental manipulations. In the class listing test, subjects recalled building names and did not use a map, and no where advantage was found. In the cued recall questionnaire, a strong where advantage was found at the one-week retention interval, but was largely reduced after six weeks. Finally, in the map test, a strong where advantage was found even after the six week interval.

These results are interpreted within the theoretical framework proposed by Kolers and Roediger (1984). To review briefly, performance in their view is described in terms of acquired skills or procedures. All memory performance, regardless of representation, reflects the degree subjects reinstate learned operations or procedures at test. Results of this research can be interpreted within this same framework, given three important specifications. First, proceduralization is operationally defined as the process of encoding and rehearsing of temporally sequenced information. Temporal sequencing simply involves learning information in a specific temporal sequence, with each sequence having a starting and

ending point. It is assumed that temporal sequencing need not involve explicit motor behavior such as walking through a campus. It is suggested that some temporal sequences are learned and practiced simply through active observation, for example, when studying a map. Further, it is proposed that such sequences can be strengthened when motor behavior is included, resulting in better retention of the sequenced information.

Second, it is suggested that spatial memories should not be thought of as simply information about the spatial relationships of objects, that is, their relative distances and directions, but must also be viewed in terms of the temporal order relationships of objects. A similar proposal was made by Healy (1978, 1982) in suggesting that spatial information was retained by means of temporal-spatial patterns. Objects in the natural environment as well as those studied two-dimensionally are easily temporally related. The same point was made by Sholl (1987) in suggesting that spatial experiences are often sequential, that is, extended over time. This is not to say that other types of information, such as, words, numbers, and letters, cannot also be related temporally. They certainly can be, but important differences may exist in the temporal sequencing of spatial versus verbal information.

Finally, it is proposed that temporal sequencing of information aides retention by providing a scheme for organizing information. Consequently, if testing conditions promote the

retrieval of this organized sequence, superior memory of the information will be demonstrated.

Evidence for procedural reinstatement can be found in both experiments of this research. The questionnaire results of Experiment 1 provide a good example of temporal sequencing involving overt motor procedures. Where information in all three testings in this experiment was better recalled than any of the other kinds of information. This supports short-term memory studies in the laboratory showing a similar advantage across time (i.e., Healy, 1982; Pezdek et al., 1986). The long-term advantage found here is attributed to the encoding and reinstatement of temporally sequenced spatial information, that is, learned procedures. Subjects in this case acquired procedures both in direct study of their schedules and by going to class.

In questioning subjects, it was revealed that when receiving a real semester schedule, students typically do two things. First, they construct a weekly calendar which indicates the periods of time when each class occurs during the school week. In other words, they took a familiar knowledge structure (the days of the week) and used it to encode the order of their classes. This structure likely included such information as times of classes, and names of courses and professors. Secondly, they carry their calendar with them each day of class until they no longer need to refer to it. By actually going to and from

classes, they develop an association between the temporal sequence of their classes and spatial relations information. In a sense, temporal sequence information provides an organizing structure for the spatial information.

The use of both temporal sequencing and spatial relations information has been demonstrated in other studies of spatial memory. Thorndyke and Hayes-Roth (1982) concluded that acquiring spatial knowledge through map study versus actual navigation through space resulted in the learning of different types of information. Map study led to survey knowledge, information about object properties, relative distances and locations. Navigation, on the other hand, led to procedural knowledge or routing information, a sequential organization of spatial knowledge. Golledge et al. (1985) in reviewing empirical studies of spatial knowledge, suggested three components of spatial knowledge could be defined: landmark knowledge (information about important objects and places), configurational knowledge (information about object relations), and route knowledge (how to get from one location to another). Unlike the first two components, route knowledge is considered procedural in nature. Evans and Pezdek (1980) demonstrated that real-life experiences on a university campus and map learning of a campus led to different mental rotation functions. They conclude that the two ways of acquiring spatial information result in different representations. These studies suggest that natural and

laboratory learning of spatial information result in acquisition of different kinds of information.

Results of Experiment 1 support the role of overt motor procedures in actual navigation. Results of Experiment 2 suggest that map study can also yield procedural learning, even when procedures involve little overt motor behavior. The where recall task in the map test and retest provides affirmative evidence. In this task, subjects had to identify all building locations on a single map. Subjects did exceptionally well at this task, showing a strong where advantage on both test and retest performance. What is being suggested is that in training subjects learned a specific temporal-spatial sequence suggested by the pattern of marks indicating building locations on the study map. Depending on the pattern, a starting and end point were selected and the locations were learned in this order. Moreover, since the when information task on the map test required recall of the temporal order of classes, the temporal-spatial pattern was further emphasized. At recall, this learned temporal-spatial sequence was reinstated, prompted by the map. Furthermore, half the subjects practiced reinstatement of this pattern on eight of the nine training trials. The other half practiced reinstatement on only one trial. Despite differences in practice, both training task groups performed equally well. These results suggest that overt motor behavior, more specifically, direct navigation, is not necessary in the

learning of temporal sequences.

The where advantage in the questionnaire test results provide further evidence for proceduralization in map study. Here again, subjects performed equally well on the where task in spite of differences in training. This evidence for temporal sequencing is particularly important because subjects in the testing task were not asked to reinstate explicitly the temporal-spatial pattern of class locations on the map, unlike subjects in the map test. Subjects instead located classes one at a time, on separate maps. Furthermore, unlike the natural learning of class locations, subjects never had the benefit of actually proceeding to and from classes. These differences might explain why the where advantage did not reach statistical significance in the questionnaire retest. The suggestion being made here is that procedures involving overt motor behavior may be better retained than those that do not. The fact that subjects who practiced retrieving the temporal-spatial pattern most in training did better at retest than those with limited practice supports this idea.

It certainly might be asked at this point whether verbal information can be temporally sequenced in the same manner being suggested for spatial information. It is very likely that subjects did sequence the what, who, where, and when information. It may have been done course-by-course and/or across courses. In other words, subjects were free to study their schedules one

course at a time and may have sequenced item information in this way or within columns, that is, by learning all course titles together. But regardless of whether verbal information was sequenced within or across courses, spatial information still showed a recall advantage.

Results of the map and class listing tests in Experiment 2 provide a good example of this difference. The map test required subjects to identify on the map the spatial locations of their classes. This task required sequencing of spatial information. Where performance was found to be better than that for what, who, or when information. The class listing test required subjects to name the buildings where classes were held. This task required subjects to proceduralize verbal information about locations. Where information recall in the class listing test did not show an advantage over what, who, or when information. All subjects in the experiment studied four types of information. Temporally sequenced information was learned and practiced in both cases and most importantly, testing required the reinstatement of the same procedures learned in training. Regardless, spatial locations were still better recalled than building names. This finding is not explainable in terms of proceduralization without modification. The proposal being put forth here is that temporally sequenced verbal information may be less easily learned, retained, and/or retrieved than is spatial information.

A number of explanations might account for this processing

disadvantage. Possibly, one of the other spatial knowledge hypotheses proposed in Chapter 3 might apply. For example, the visual code hypothesis suggested that visual information might be important in the encoding and/or retrieval of spatial knowledge. Numerous studies involving spatial memory have suggested the role of visual imagery (Kosslyn, 1976; Paivio, 1971, 1986). McNamara (1986) concluded that spatial knowledge is stored in both image-like networks and propositionally, emphasizing the redundancy in representation. McNamara also pointed out that an image-like representation is best for encoding spatial configurations while a propositional format is best for semantic or logic knowledge. Presson and Hazelrigg (1984) distinguished between primary and secondary spatial knowledge. Primary knowledge is acquired by actual experience in a real environment. Secondary knowledge is the type acquired through maps. They suggested that the secondary spatial knowledge contains picture-like properties, sensitive to orientation and alignment. For much the same reason, the spatial retention hypothesis might also explain this difference. Perhaps spatial information is uniquely retained in memory simply because it is visually represented. Bahrick et al. (1975) have suggested that visual information is retained much longer in memory than verbal information. Another possibility is that spatial information is automatically encoded, thus making sequencing easier. Such a proposal has been made by Mandler, Seegmiller and Day (1977) and

Hasher and Zacks (1979), but has not gone without considerable criticism (see Naveh-Benjamin, 1987). Finally, it could be argued that people simply have more experience at sequencing spatial information than verbal information. Retrieval of temporally sequenced spatial information is a daily event for most people.

In sum, two factors are considered important in explaining the spatial information advantage in this research. First, results particularly from the natural learning of course schedules in Experiment 1 and the map test results of Experiment 2 suggest that temporal sequencing of spatial information and the reinstatement of this sequence at recall provided a performance advantage for where information. Secondly, the results of the map and class listing tests in Experiment 2 suggest that temporally sequenced verbal information shows a processing disadvantage compared to spatial information, even when the testing task promotes the reinstatement of the same sequences learned and practiced in training.

It is important to realize that the proposed procedural account of the where advantage is a tentative one. As previously discussed, several other explanations for the where advantage are possible. One of the stronger alternative accounts suggests the map played an important role in the where task by making the task a recognition test versus a recall test, as with the other types of information. Furthermore, it has been assumed that the where

advantage in Experiments 1 and 2 can be accounted for using the same explanation; the temporal sequencing of spatial information. This does not, however, have to be the case. Training in each experiment was substantially different, even though the laboratory training in Experiment 2 was designed to be similar to the natural training of Experiment 1. It is certainly possible that some other learning factor not yet considered was operative in Experiment 1. Finally, it must be acknowledged that both experiments in this research provide only indirect evidence for the role of temporal sequencing in the recall of spatial information. Using the results of this research as a starting point, a more direct test of the importance of temporal sequencing in the where advantage needs to be conducted. Thus, several modifications of the current experiments are proposed.

The importance of temporal sequencing in the spatial information advantage can be further explored by restricting the order in which subjects recall building locations on the map. If subjects continue to show a where advantage after learning spatial information in one order and then recalling it in another order, the temporal sequence hypothesis would not be supported. To compare better the sequencing of verbal and spatial information, another task modification might include using the campus map to recall building names, as was done for building locations. Subjects could be asked to identify location by simply marking the building with an "X" or by marking the

building with its name. In addition recall performance could be compared with the temporal order in which the building names were presented for study and with the order suggested by the locations of buildings on the map. The task could be further modified by having subjects learn building names directly from the map, that is, by marking each building with its name. This change provides the extra advantage of examining the role of the campus map as an additional recall cue. In this case both spatial and verbal information recall tasks would use the map.

The major focus in this discussion to this point has been on the advantage for spatial information over temporal or item information in recall. Before concluding, it is important to review results in terms of the second goal of this research, the comparison of natural and laboratory studies of human memory.

Studies of naturally learned memories are typically plagued with control difficulties. The three common methodologies, the probe, diary, and questionnaire methods, each have their unique problems. This study of natural memory was designed to make use of the best aspects of each method while avoiding their limitations. Consequently, all events studied were pre-established, yet none were pre-recorded by subjects. It used prompt words, but each cue was specific to only one event. Furthermore, it tested memories over specified retention periods rather than having subjects recall memories from an unspecified period in the past. Finally, it examined performance across a

large number of subjects, not being limited to a single subject design. It is not being suggested that this study controlled all aspects of the learning situation. Such factors as degree of original learning and the amount of subsequent rehearsal were not controlled, only estimated. What is being suggested is that the methodological approach used in this study produced reliable and consistent results. These results therefore, lend validity to a spatial, temporal, and item distinction in memory and make possible important comparisons with other natural studies and related laboratory research.

The most obvious comparison is between these results and Wagenaar's (1986). As noted earlier, Wagenaar did not compare recall across the what, who, where, and when aspects of his events. It was obvious from his data, however, that the when information (dates) showed the poorest recall performance. In this research, the same relatively poor recall of when information (class times) was found, but only in Experiment 2 where class times were learned without actually going to classes. Results of the questionnaire test and retest and the class listing retest showed poorest recall of when information (class times) relative to what, who, or where information. This was true despite the fact that the training data suggested when information was learned rapidly, and was the only type of information in which performance reached 100% correct, remaining so for the last three training trials. It is also important to

point out that when information recall in the map test and retest did not show the poorest recall. Who information showed the poorest performance. In the map test and retest subjects recalled the relative order of their classes, not class times. Comparing the class listing test and retest results with the map test and retest suggests that subjects performed better at recalling class order than class start times. These results are consistent with other natural memory studies emphasizing the unique role of temporal information (i.e., Loftus & Marburger, 1983; Baddeley, Lewis, & Nimmo-Smith, 1978) and its possibly unique rate of forgetting (i.e., Thompson, 1982). A second area of similarity involves cuing results. Wagenaar found that "what the event was" served as the best cue for recall and "when the event occurred" served as the worst cue. In this study it was found that what and who information served equally well as cues and that where and when were equally bad. This difference in results is likely due to differences in the type of event recalled, that is, daily experiences versus class schedules.

Probably the most significant comparison that can be made between this research and other natural studies has to do with the reported retention rates of natural memories. For the most part, natural memory studies report impressively high recall rates. Linton (1978), for example, in a six year diary study found that at the end of six years, she remembered 68% of the material. Barclay and Wellman (1986) suggest that natural

studies often overestimate recall accuracy because of such problems as recording bias. In this study, while percentages between 60 and 80 percent one year after training map may look impressive, the 16 week-long rehearsal period (the semester) and the potential for periodic reminding after the semester ended, must be kept in mind. Furthermore, performance dropped to between 40 and 50 percent after the first year (except for where information). This result brings up a second important point related to retention rates. There is some controversy in the natural memory literature about the shape of the retention function for natural memories. Those using the probe method (Rubin, 1982; Rubin, Wetzler, & Nebes, 1986) contend that natural and laboratory memories follow the same retention function, that is, the classic forgetting curve with an initial rapid decline followed by gradual slowing. Other researchers using the diary method have provided evidence for a slower rate of forgetting (Linton, 1978, 1982; Wagenaar, 1986). Results of this study clearly demonstrated the classic forgetting function, with the greatest loss of information occurring between years one and two (except for where information) and a much slower loss between years two and three. These results are comparable with Bahrick's (1984) study of very long-term memory for Spanish. In Bahrick's results, information continued to be loss until the fifth or sixth year, after which the function flattened out, showing evidence, in his terms, for storage of information in permastore.

It is interesting to speculate how much students in this study might have "permanently" stored about their class schedules. This question suggests yet another future direction for this project.

In conclusion, this research found a consistent advantage for spatial information (class locations) over temporal information (class times), and item information (course titles and professor names). Results are interpreted in terms of Kolers and Roediger (1984) procedural view of memory performance. More specifically, it is tentatively proposed that both spatial and item information can be organized into temporal sequences (proceduralized), but that spatial information has an advantage either in the encoding, retention, or retrieval of these sequences, thus showing a recall advantage. The specific nature of this sequencing advantage is yet to be determined. Only by examining spatial memories learned both naturally and in the laboratory were these findings fully realized.

REFERENCES

Anderson, J. R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.

Anderson, J. R. (1983). The architecture of cognition. Cambrige, MA: Harvard University Press.

Baddeley, A. D., Lewis, V., Nimmo-Smith, I. (1978). When did you last . . . ? In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), Practical aspects of memory, (pp. 77-83). New York: Academic Press.

Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. Journal of Experimental Psychology: General, 113, 1-29.

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. Journal of Experimental Psychology: General, 104, 54-75.

Barclay, C. R., & Wellman, H. M. (1986). Accuracies and inaccuracies in autobiographical memory. Journal of Memory and Language, 25, 93-103.

Brown, R., & Kulik, J. (1977). Flashbulb memories. Cognition, 5, 73-99.

Cohen, N. J. (1984). Preserved learning capacity in amnesia: Evidence for multiple memory systems. In L. R. Squire, & N. Butters (Eds.), Neuropsychology of memory (pp. 83-103). New York: The Guilford Press.

Crovitz, H. F., & Schiffman, H. (1974). Frequency of episodic memories from early childhood by years of age. Bulletin of the Psychonomic Society, 4, 517-518.

Dean, R. S., & Kulhavy, R. W. (1981). Influence of spatial organization in prose learning. Journal of Educational Psychology, 73, 57-64.

Evans, G. W., & Pezdek, K. (1980). Cognitive mapping: Knowledge of real-world distance and location information.

Journal of Experimental Psychology: HumanLearning & Performance, 6, 13-24.

Fitts, P. M., & Posner, M. I. (1967). Human Performance. Belmont, CA: Brooks/Cole.

Franklin, H. C., & Holding, D. H. (1977). Personal memories at different ages. Quarterly Journal of Experimental Psychology, 29, 527-532.

Golledge, R. G., Smith, T. R., Pellegrino, J. W., Doherty, S., & Marshall, S. P. (1985). A conceptual model and empirical analysis of children's acquisition of spatial knowledge. Journal of Environmental Psychology, 5, 125-152.

Glenberg, A. M. (1979). Component-levels theory of the effects of spacing repetitions on recall and recognition. Memory & Cognition, 7, 95-112.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. Journal of Experimental Psychology: General, 108, 356-388.

Healy, A. F. (1974). Separating item from order information in short-term memory. Journal of Verbal Learning and Verbal Behavior, 13, 644-655.

Healy, A. F. (1978). A Markov model for short-term retention of spatial location information. Journal of Verbal Learning and Verbal Behavior, 17, 295-308.

Healy, A. F. (1982). Short-term memory for order information. In G. H. Bower (Ed.), The Psychology of Learning and Motivation, (Vol. 16, pp. 191-238). New York: Academic Press.

Herrmann, D. J., & Neisser, U. (1978). A inventory of everyday memory experiences. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), Practical aspects of memory (pp. 35-51). New York: Acdemic Press.

Kosslyn, S. M. (1976). Can imagery be distinguished from other forms of interval representation? Evidence from studies of information retrieval times. Memory & Cognition, 4, 291-297.

Kosslyn, S. M. (1987). Seeing and imaging in the cerebral hemispheres: A computational approach.

Psychological Review, 94, 148-175.

Kolers, P. A., & Roediger, H. L., III. (1984). Procedures of mind. Journal of Verbal Learning and Verbal Behavior, 23, 425-449.

Lee, C. L., & Estes, W. K. (1981). Item and Order information in short-term memory: Evidence for multilevel perturbation processes. Journal of Experimental Psychology: Human Learning and Memory, 7, 149-169.

Linton, M. (1978). Real world memory after six years: An in vivo study of very long term memory. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), Practical aspects of memory (pp. 69-76). New York: Acdemic Press.

Linton, M. (1982). Transformations of memory in everyday life. In U. Neisser (Ed.), Memory observed: Remembering in natural contexts (pp. 77-91). San Francisco: W. H. Freeman.

Loftus, E. F., & Marburger, W. (1983). Since the eruption of Mount St. Helens, has anyone beaten you up? Improving the accuracy retrospective reports with landmark events. Memory & Cognition, 11, 114-120.

Mandler, Seegmiller, & Day, (1977). On the coding of spatial information. Memory & Cognition, 5, 10-16.

McNamara, T. P. (1986). Mental representations of spatial relationships. Cognitive Psychology, 18, 87-121.

Murdock, B. B. (1976). Item and order information in short-term serial memory. Journal of Experimental Psychology: General, 105, 191-216.

Naveh-Benjamin, M. (1987). Coding of spatial location information: An automatic process? Journal of Experimental Psychology: Learning, Memory and Cognition, 4, 595-605.

Paivio, A. (1971). Imagery and verbal processes. New York: Holt, Rinehart, & Winston.

Paivio, A. (1986). Mental representations: A dual coding approach. New York: Oxford University Press.

Pezdek, K., Roman, Z., Sobolik, K. G. (1986). Spatial memory for objects and words. Journal of Experimental Psychology: Learning, Memory, and Cognition, 12, 530-537.

Pillemer, D. B. (1984). Flashbulb memories of the assassination attempt of President Reagan. Cognition, 16, 63-80.

Presson, C. C., & Hazelrigg, M. D. (1984). Building spatial representations through primary and secondary learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 716-722.

Robinson, J. A. (1976). Sampling autobiographical memory. Cognitive Psychology, 8, 578-595.

Rubin, D. C. (1982). On the retention function for autobiographical memory. Journal of Verbal Learning and Verbal Behavior, 21, 21-38.

Rubin, D. C., Wetzler, S. E., & Nebes, R. D. (1986). Autobiographical memory across theh lifespan. In D. C. Rubin (Ed.), Autobiographical memory (pp. 202-221). Cambridge: Cambridge University Press.

Salthouse, T. A. (1974). Using selective interference to investigate spatial memory representations. Memory & Cognition, 2, 749-757.

Salthouse, T. A. (1975). Simultaneous processing of verbal and spatial information. Memory & Cognition, 3, 221-225.

Schwartz, N. H., & Kulhavy, R. W. (1981). Map features and the recall of discourse. Contemporary Educational Psychology, 6, 151-158.

Shiffrin, R. M., & Cook, J. R. (1978). Short-term forgetting of item and order information. Journal of Verbal Learning and Verbal Behavior, 17, 189-218.

Sholl, M. J. (1987). Cognitive maps as orienting schemata. Journal of Experimental Psychology: Learning, Memory, & Cognition, 13, 615-628.

Thorndyke, P. W., & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired from maps and navigation.

Cognitive Psychology, 14, 560-589.

Thompson, C. P. (1982). Memory for unique personal events: The roommate study. Memory & Cognition, 10, 324-332.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of memory. New York: Academic Presss.

Tulving, E. (1985). How many memory systems are there? American Psychologist, 40, 385-398.

Wagenaar, W. A. (1978). Recalling messages broadcast to the general public. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), Practical aspects of memory (pp. 128-136). New York: Acdemic Press.

Wagenaar, W. A. (1986). My memory: A study of autobiographical memory over six years. Cognitive Psychology, 18, 225-252.

White, R. T. (1982). Memory for personal events. Human Learning, 1, 171-183.

Yarmey, A. D., & Bull, M. P., III. (1978). Where were you when President Kennedy was assassinated? Bulletin of the Psychonomic Society, 11, 133-135.

APPENDIX A

Sample of Questionnaire

Course 1

1. The name of the course was______________________________.

Complete the following questions as best you can.

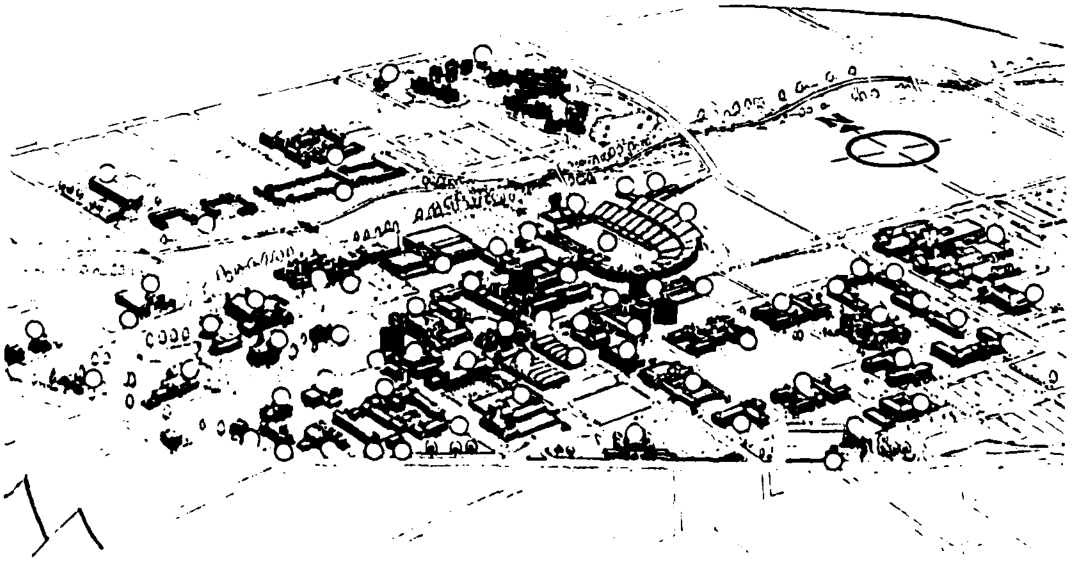
2. Where was the course held? (mark the map below)

3. When was the course held?

(give the exact time of the day)______________________.

4. Who was the instructor?_________________________________.

(Note: the map below has been reduced in size)



APPENDIX B

Sample Questions from Questionnaire, Experiment 1

1. Rate how much you enjoyed this course.
   1 2 3 4 5 6 7
   Very little ... Very much

2. Rate how much you enjoyed the instructor.
   1 2 3 4 5 6 7
   Very little ... Very much

3. Rate how much you enjoyed the time when this course was held.
   1 2 3 4 5 6 7
   Very little ... Very much

4. Rate how much you enjoyed the building where this course was held.
   1 2 3 4 5 6 7
   Very little ... Very much

5. Was this course within your major field of study?
   (circle one) Yes No

6. How many times have you taken a course within the same department, i.e., in the same subject area, since this course ended?
   (circle one) 0 1 2 3 4 or more times

7. How many times have you had this instructor since this course ended?
   (circle one) 0 1 2 3 4 or more times

8. Have you had (or do you currently have) this instructor as an advisor?
   (circle one) Yes No

9. How many time have you had a course at this exact time of the day since this course ended?
   (circle one) 0 1 2 3 4 or more times

10. How many courses have you had in this same building since this course ended?

(circle one) 0 1 2 3 4 or more times

11. Rate how much you enjoyed this entire semester

 1 2 3 4 5 6 7

 Very little Very much

APPENDIX C

Sample Class Schedule

Dept Abbev.	Course Number	Course Title	Instructor	Class Time	Days	Location Bldg	Room
PHIL	100	INTRO TO PHILOSOPHY	ROGERS	1230-1345	T TH	GUGG	201
P E	113	CONDITIONING	FOWLER	1000-1050	M W F	CGYM	E012
A S	131	VISUAL ARTS/HUMAN EXP	BERNIER	1100-1215	T TH	F A	N185
ENGL	190	INTRO TO SHAKESPEARE	WALL	900-950	M W F	EDUC	143
R ST	260	WORLD RELIGIONS-WEST	CHERNUS	1400-1515	T TH	HALE	103

APPENDIX D

Sample Class Listing Task and Test Form

INSTRUCTIONS: Complete the missing information in the following class listing.

Dept Abbev.	Course Number	Course Title	Instructor	Class Time	Building
R ST	260				
CHEM	103				
A S	130				
APAS	112				
A M	135				

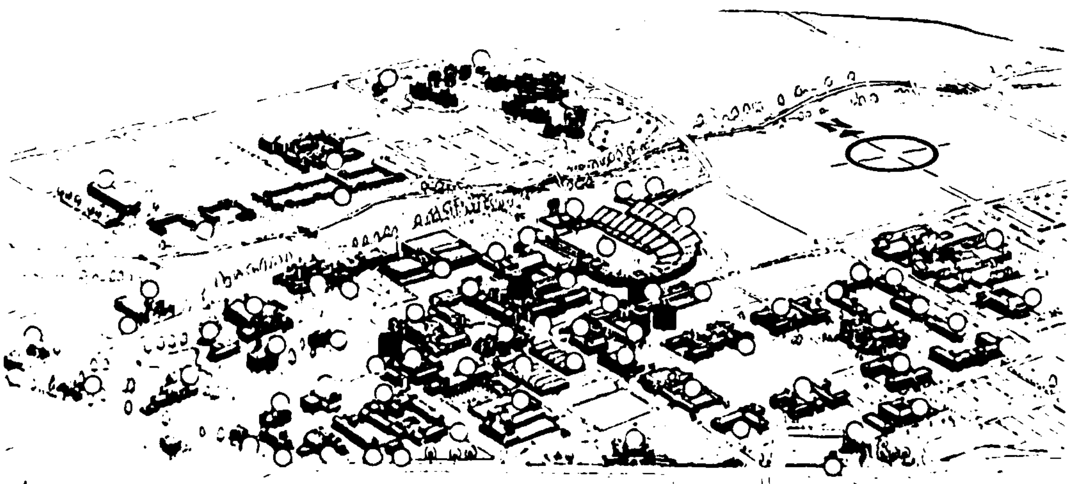
APPENDIX E

Sample Map Task and Test Form

INSTRUCTIONS: Complete the missing information in the following class listing. In the class order column indicate the order in which your classes took place. Mark "1" next to the course which occurred first during the school week, a "2" next to the course occurring second, and so on. The school week begins on Monday morning and ends Friday afternoon. Consider only the first time a class occurred during a week.

Dept Abbev.	Course Number	Course Title	Instructor	Class Order
PHIL	100			
P E	113			
A S	131			
ENGL	190			
R ST	260			

INSTRUCTIONS: Mark on the map below the locations of your classes shown on your class listing. Color in the blank circle next to the building where your class is held.



APPENDIX F

Sample Questions from Questionnaire, Experiment 2

1. Describe how you learn your own university class schedule at the start of a semester.

2. How was the way you studied the fictitious schedule in this experiment the same or different from how you study your own schedule in real-life.

3. How many times did you study and then recall the fictitious schedule in this experiment?__________.

4. When learning the fictitious schedule in this experiment, what did you find most important to study?

5. When learning the fictitious schedule in this experiment, what did you find least important to study?

6. Between sessions of this experiment, did you find yourself thinking about your fictitious schedule?___________. If you did, how much did you think about it? (rate how much on a scale from 1 to 7, with "1" being "very little" and "7" being "very much")_________.

APPENDIX G

Questionnaire Test and Retest Error Analysis

Table 22 provides mean error rates for test data across information types by training, task, and number of courses.

Table 22

Mean Error Rates for Questionnaire Test Data Across Information Types as a function of Training, Task, and Number of Courses

	Information type *				
	what	who	where	when	
Training *					
Massed	1.37	1.17	0.58	1.25	(1.09)
Spaced	0.42	0.38	0.46	0.33	(0.40)
Task					
Map	1.04	0.96	0.54	0.92	(0.86)
Class Listing	0.75	0.58	0.50	0.67	(0.63)
Courses					
4	0.88	0.71	0.58	0.79	(0.74)
5	0.92	0.83	0.46	0.79	(0.75)
	(0.90)	(0.77)	(0.52)	(0.79)	(0.74)

Note. Number of subjects per cell was 24. Column and row means are provided in parentheses.

* Significant main effect at $p < .05$.

Consistent with the results from the accuracy data was a highly significant main effect of training type [$F(1,40) = 22.29$, $MSe = 23.38$]. Subjects in the massed training condition scored more errors than those in the spaced training condition (Ms = 1.09 and 0.40, respectively). There was no significant main effect of task, though, like the accuracy data, there was a tendency for those assigned to the map task to make more errors (Ms = 0.86 for the map task and .63 for the class listing task). There was no main effect of number of courses. There was a significant Number of Courses X Task interaction, however [$F(1,40) = 4.18$, $MSe = 4.38$]. Following the accuracy data, there was no difference in the number of errors across tasks when there were 4 courses in the schedule (Ms = 0.71 for the map task and 0.77 for the class listing task). With 5 courses though, those assigned to the class listing task had fewer errors than those in the map task (Ms = 1.02 for the map task and 0.48 for the class listing task).

Also consistent with results for subject accuracy, the main effect of information type on the questionnaire test was found to be significant [$F(3,120) = 3.20$, $MSe = 1.21$]. Subjects made fewer errors for where information than for either of the other three types of information. The mean error rates for what, who, where, and when, in order, were 0.90, 0.77, 0.52, and 0.79. This difference was clearly significant for what versus where information [$F(1,40) = 8.44$, $MSe = 6.75$] and who versus where [$F(1,40) = 5.90$, $MSe = 3.00$], and marginally significant for when

versus where [$F(1,40) = 3.82$, $MSe = 3.52$, $p < .06$]. Furthermore, as in the accuracy data, training type interacted significantly with information type [$F(3,120) = 4.77$, $MSe = 1.81$] (see Figure 10).

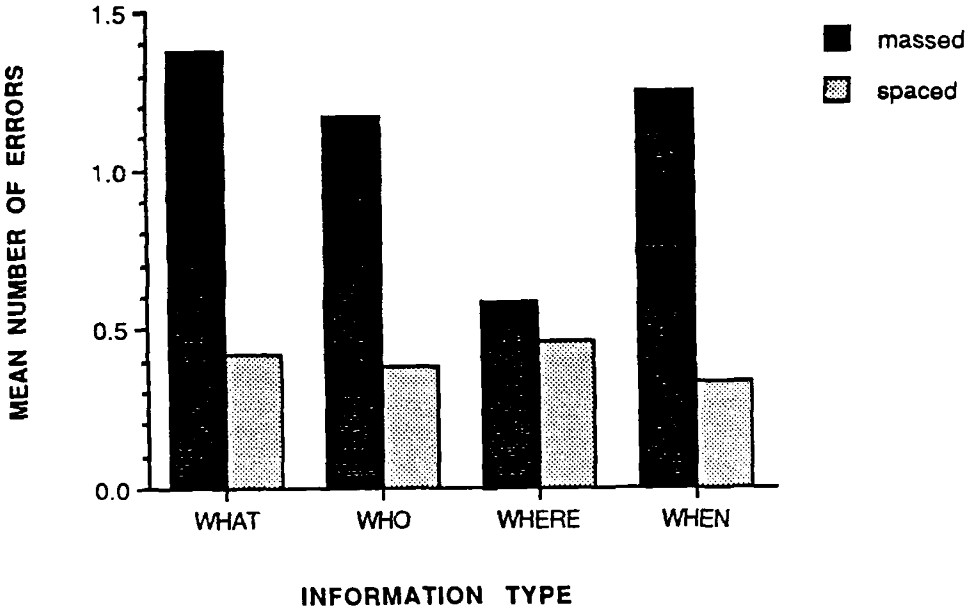


Figure 10. Information Type X Training Type interaction in the questionnaire test error data, Experiment 2.

Subjects receiving massed training scored more errors than spaced training subjects for what, who, and when information. The main effect of training for these three types of information was highly significant, for what [$F(1,40) = 17.06$, $MSe = 11.02$], for who [$F(1,40) = 13.78$, $MSe = 7.52$], and for when [$F(1,40) = 22.00$, $MSe = 10.83$]. This was not true for where information. In this case, the effect of training type was not significant, likely reflecting a floor effect for where information.

Results of the error analysis for questionnaire retest data generally followed the same pattern just described for the test data. Mean error rates across information types by training, task, and number of courses are provided in Table 23.

Table 23

Mean Error Rates for Questionnaire Retest Data Across Information Types as a function of Training, Task, and Number of Courses

	Information type				
	what	who	where	when	
Training *					
Massed (n = 19)	1.42	1.26	1.05	1.47	(1.30)
Spaced (n = 17)	1.00	0.76	0.65	0.88	(0.82)
Task					
Map (n = 15)	1.00	1.00	0.53	1.27	(0.95)
Class Listing (n = 21)	1.38	1.05	1.10	1.14	(1.17)
Courses					
4 (n = 18)	1.44	0.89	0.89	1.17	(1.10)
5 (n = 18)	1.00	1.17	0.83	1.22	(1.06)
	(1.22)	(1.30)	(0.86)	(1.19)	(1.08)

Note. Column and row means are provided in parentheses.

* Significant main effect at $p < .05$.

The only between-subjects factor to reach statistical significance was training type type [$F(1,28) = 8.94$, $MSe = 7.70$]. Massed training subjects made more errors ($M = 1.30$, $n = 19$) than spaced training subjects ($M = .82$, $n = 17$). No interactions were significant. Furthermore, there was no significant main effect of information type. Mean error rates for what, who, where, and when, respectively, were 1.22, 1.03, .86, 1.19. Despite the lack of overall significance, single degree of freedom comparisons showed the mean number of errors for where information to be significantly below the mean for when [$F(1,28) = 5.80$, $MSe = 5.21$] and marginally below the mean for what information [$F(1,28) = 3.09$, $MSe = 3.68$, $p < .09$]. This is in contrast to the relatively close performance for what, who, and where information found for accuracy on the questionnaire retest (Ms = 60.19%, 62.28%, 64.36%, respectively). These results suggest that the lack of clear superior performance for where information on the retest might be masked by differences in the types of errors made from test to retest. These differences are examined below.

Additional analyses were performed comparing directly errors on the questionnaire test with errors on the retest. Only the 36 subjects who participated in both the test and retest were used in this analysis. All four types of information showed a significant increase in errors from test to retest (see Table 24).

Table 24

Mean Number of Errors by Information Type for Questionnaire Test and Retest

	Information type				
	what	who	where	when	
Test	0.89	0.81	0.44	0.81	(0.74)
Retest	1.22	1.03	0.86	1.19	(1.08)

Note. Row means are provided in parentheses. Number of subjects for both test and retest was 36.

* Significant differences from test to retest, $p < .05$.

The largest increase occurred for where information with means of 0.44 and 0.86 for test and retest, respectively [$F(1,28) = 9.22$, $MSe = 5.32$]. This was followed by when information [$F(1,28) = 5.43$, $MSe = 5.78$, $Ms = 0.81$ and 1.19 for test and retest, respectively], what information [$F(1,28) = 5.46$, $MSe = 2.85$, $Ms = 0.89$ and 1.22], and finally who information [$F(1,28) = 4.36$, $MSe = 1.92$, $Ms = 0.81$ and 1.03, respectively].

END